



TITLE:

Studies on Diversity and Stress Responses of
the Cytochrome P450 Monooxygenase
System in Higher Plants(Dissertation_全文)

AUTHOR(S):

Mizutani, Masaharu

CITATION:

Mizutani, Masaharu. Studies on Diversity and Stress Responses of the Cytochrome P450 Monooxygenase System in Higher Plants. 京都大学, 1997, 博士(農学)

ISSUE DATE:

1997-09-24

URL:

<https://doi.org/10.11501/3130714>

RIGHT:

Studies on Diversity and Stress Responses of the Cytochrome P450 Monooxygenase System in Higher Plants

Masaharu Mizutani

1997

.....8*

..... 90

..... 97

..... 98

Contents

General Introduction	----- 1
Chapter I:	----- 7
Purification and Characterization of a Cytochrome P450 (Cinnamate 4-Hydroxylase) from Etiolated Mung Bean Seedlings	
Chapter II:	----- 21
Molecular Cloning and Sequencing of a cDNA Encoding Mung Bean Cytochrome P450 (P450C4H) Possessing Cinnamate 4-Hydroxylase Activity	
Chapter III:	----- 28
Isolation of a cDNA and a Genomic Clone Encoding Cinnamate 4-Hydroxylase from <i>Arabidopsis thaliana</i> and its Expression Manner in <i>Planta</i>	
Chapter IV:	----- 46
Two Isoforms of NADPH-Cytochrome P450 Reductase in <i>Arabidopsis thaliana</i> : Gene Structure, Heterologous Expression in Insect Cells, and Differential Expression in <i>Planta</i> .	
Chapter V:	----- 68
Cytochrome P450 Superfamily in <i>Arabidopsis thaliana</i> : cDNAs Isolation, Differential Expression, and RFLP Mapping of Multiple Cytochromes P450	
Conclusion	----- 88
References	----- 90
List of Publications	----- 97
Acknowledgments	----- 98

General Introduction

Biological oxidation of organic substances is an important process to provide physiologically active compounds which are used in various biochemical processes. Biological oxidation in general proceeds *via* two different mechanisms. One mechanism in which the oxygen atom is derived from a water molecule is catalyzed by dehydrogenases, which are involved in hydration and hydrolysis of a substrate. The other mechanism is based on the activation of molecular oxygen and is catalyzed by two classes of enzymes, oxidases and oxygenases. Oxidases transfer electrons from a substrate to oxygen, thereby oxidizing the substrate. On the other hand, oxygenases transfer oxygen to a substrate *via* reductive cleavage of molecular oxygen and insertion of oxygen into the substrate. Cytochromes P450 belong to the latter class of the molecular activating enzymes and are involved in biological oxidation reactions in a variety of metabolic processes of endogenous and exogenous substrates.

Cytochrome P450 monooxygenase system

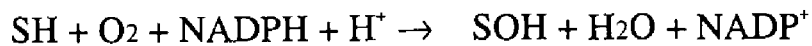
a. General properties

Cytochrome P450 (P450) is the generic name for a large family of a *b*-type haemoprotein and contains iron protoporphyrin IX as a prosthetic group. The term 'cytochrome P450' has been proposed by Omura and Sato (1964) for a haemoprotein with an unusual absorption maximum of its ferrous carbonyl complex. Thus, a cytochrome P450 enzyme is defined by its absorption spectrum which exhibit the Soret absorption maximum at 450 nm by the binding of CO to the reduced enzyme. This atypical spectral property arises from the coordination of a thiolate anion ($-S^-$, derived from conserved cysteinyl residue of the protein) to the heme iron as the fifth ligand. P450s are widely distributed in the biological kingdom; in bacteria they are soluble proteins, whereas in eukaryotic cells they are bound to the microsomal membrane (in some cases, also to the inner mitochondrial membrane).

b. Catalytic activity

P450s are monooxygenases that catalyze the incorporation of one oxygen atom from molecular oxygen into lipophilic, low-molecular-weight organic compounds with the

aid of reducing equivalents (electrons) derived from NADPH (NADH in bacteria). The reaction can be expressed as follows:



In detail, P450s introduce an oxygen atom into a substrate (SH) by the reaction cycle as shown in Fig. 1. The P450 monooxygenation reaction is initiated by binding a substrate to the oxidized state of P450. The first electron is transferred from the cytosolic NADPH to the P450-SH complex via a specific electron-transporting protein (as described below), resulting in conversion of ferrous heme iron to the ferric form. Molecular oxygen binds the sixth heme ligand and is converted to superoxide anion (O_2^-). After insertion of the second electron, the oxygen species is further converted to a ferric iron coordinated peroxide complex ($\text{Fe}^{3+} - \text{O}_2^{2-}$). The complex reacts with H^+ and releases water, and an activated oxygen molecule remains at the sixth heme ligand of P450. The activated oxygen molecule then reacts with the substrate, and the hydroxylated product is liberated and the ferric form of P450 is ready to enter the next reaction cycle.

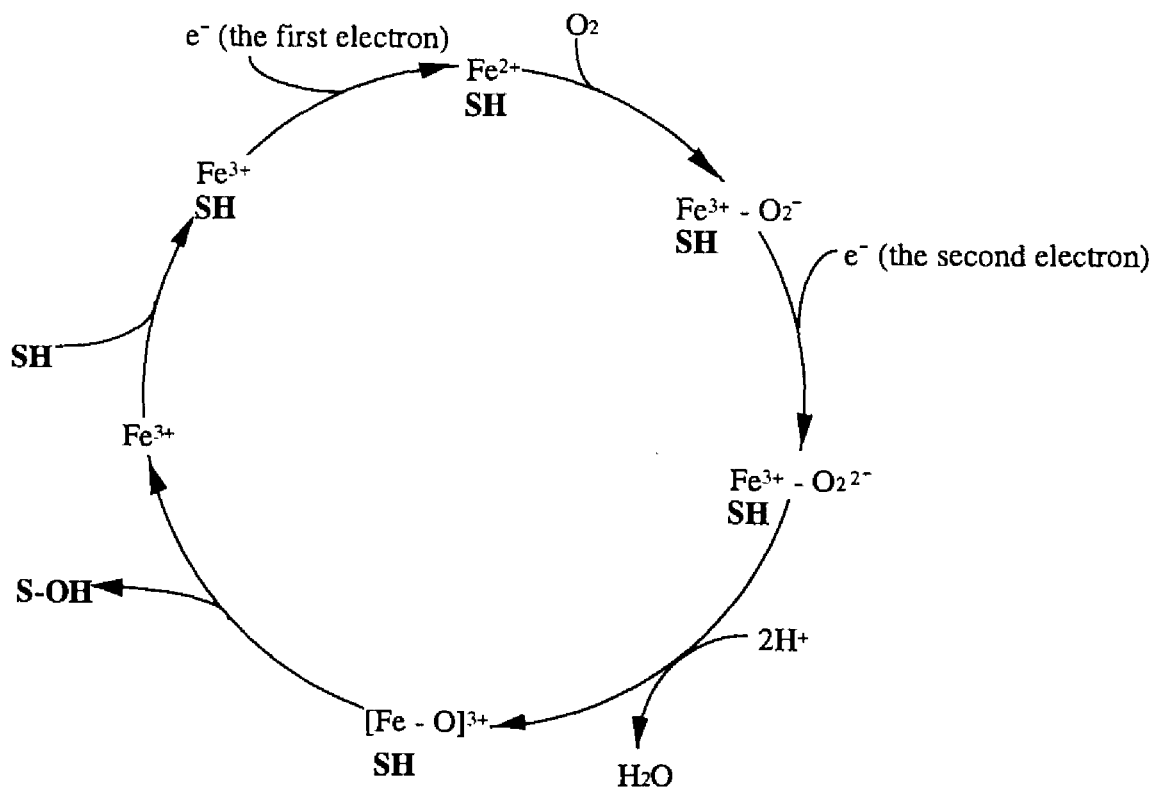
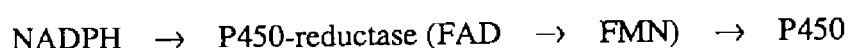


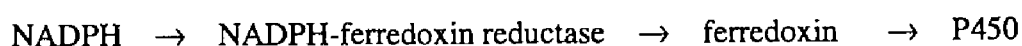
Figure 1. Reaction cycle of cytochromes P450.
 Fe^{3+} , the oxidized form of P450; Fe^{2+} , the reduced form of P450; SH, substrate

c. Electron transport scheme

In microsomes the transfer of electrons from NADPH to P450s is mediated by the flavoprotein NADPH-cytochrome P450 reductase (P450-reductase). P450-reductase is an essential component of P450 monooxygenase system. P450-reductase is a unique flavoprotein which is bound to ER membrane and contains both FMN and FAD as prosthetic groups, and its molecular weight is estimated as 74 - 80 kDa. It functions in electron transfer from NADPH to P450s as the following route:



On the other hand, in mitochondria and bacteria, two proteins, ferredoxin and NADPH-ferredoxin reductase, mediate electron transfer from NADPH (NADH in bacteria) to P450s. NADPH-ferredoxin reductase has a molecular weight in the range from 45 - 52 kDa and contains one mole of FAD. Ferredoxin contains a two-iron, two-sulfur prosthetic group and has a molecular weight of 14 kDa. P450s receive electrons from these enzymes as the following route:



In some cases, P450 reactions are stimulated by addition of cytochrome *b5*, suggesting the participation of cytochrome *b5* in the electron transport scheme of P450s.

d. Functions

As monooxygenases, many P450s are involved in the metabolism of endogenous lipophilic compounds: *e.g.* biosynthesis of sterols, steroid hormones, bile acids, prostaglandins, *etc.*; activation of vitamin D₃; fatty acid ω -oxidation. The second metabolic role of eukaryotic P450s is oxidative detoxification of a large variety of xenobiotics such as drugs, natural plant products, food additives, agrochemicals, and environmental pollutants.

e. Multiplicity

At least in mammals one organism has a large number of P450 genes (several hundreds?), but the actual number of P450 genes is not yet known. Some of these genes are expressed constitutively and in a tissue-specific manner. The expression of the rest of P450 genes is induced by treatments that alter physiological and pharmacological

conditions of the organism. For instance, administration of drugs, or other xenobiotics to animals leads to the induction of several species of P450s in liver, lung, stomach, *etc.* depending on the xenobiotics given.

f. Substrate specificity

Generally P450s involved in endogenous metabolism possess more or less strict substrate specificities as in the case of most other enzymes. The substrate specificities of xenobiotic-metabolizing P450s are, however, unusually broad and overlapping. Such P450s can act on a large variety of organic compounds (over a thousand?) that are completely unrelated structurally. These broad specificities of P450s, together with their multiplicity, give the organism the ability to cope with a wide variety of xenobiotics that can be ingested. One of the major questions in P450 research has been to understand the basis for this broad substrate specificity.

g. P450 superfamily

Since the cloning of the first animal P450 genes in 1982, to date over 400 P450 genes from microbial, yeast, mammals, and recently, plants have been sequenced. The amino acid sequences of many P450s have been compared and classified on the basis of proposed evolutionary relationships of the corresponding genes as described below. First, a P450 protein sequence from one gene family is defined as usually having less than 40% resemble to that from any other family. Second, within the same subfamily mammalian sequences are always >55% identical, and inclusion of more distant species (*e.g.*, mammals and fishes) within the same subfamily drops this value to >46%.

Cytochromes P450 in higher plants

Cytochrome P450-like pigments were first recognized in microsomes from a variety of plant species more than twenty years ago (Markham et al, 1972) and the properties of the plant cytochromes were found to be very similar to those of mammalian P450s (Rich and Bendall, 1975). Since then, it has been identified by biochemical approaches that various plant P450s are involved in many oxidation reactions of secondary metabolism as monooxygenases (West, 1980).

In mammalian cells, P450s are found in ER and mitochondrial membranes, whereas in bacteria P450s are soluble. In higher plants, most of P450s are usually found in microsomal fractions: *e.g.* cinnamate 4-hydroxylase; N-demethylase; kaurene oxidase, *etc.* (cited in West, 1980). However, there are several plant P450s which are localized in other than microsomes: monoterpene hydroxylase localized in provacuolar membranes (Madyasha et al., 1977); allene oxide synthase having a N-terminal chloroplast transit peptide (Song et al., 1993); a P450 specifically localized in guayule rubber particles (Pan et al., 1995); and benzoic acid 2-hydroxylase as a soluble form (Leon et al., 1995).

In higher plants, P450s play crucial roles in biosynthesis of a variety of endogenous lipophilic compounds such as fatty acids, sterols, phenylpropanoids, terpenoids, phytoalexins, and gibberellins (Donaldson and Luster 1991; Bolwell et al., 1994). In addition, oxidative detoxification of a number of herbicides in plant tissues is also achieved by the P450-dependent monooxygenase system (Riviere and Cabbane 1987; Donaldson and Luster 1991; Hatzios 1991; Sandermann 1992). Thus, many plant P450s are involved in various metabolic processes and each P450 should be differentially expressed in response to developmental and environmental cues. Several members of plant P450s are known to be induced by environmental stresses. The activity of cinnamate 4-hydroxylase was found to be induced by light and wounding treatments (Benveniste et al., 1977; Lamb 1977; Oba and Conn, 1988). Pathogen attack was known to induce phytoalexin synthesis and the fungal elicitor also induced the activities of two P450s, cinnamate 4-hydroxylase and 3,9-dihydroxypterocarpan 6 α -hydroxylase in soybean (Kochs and Grisebach, 1989). P450s are also induced in response to chemical stresses. Plant exposure to divalent cations such as Mn²⁺ or Cd²⁺ has been shown to increase the P450 specific content and related activities in artichoke (Reichhart et al., 1980). The organic chemicals which induce P450 expression in mammals (*e.g.* phenobarbital, clofibrate, ethanol, *etc.*) also induced plant P450s (Reichhart et al., 1980; Zimmerlin and Durst, 1992). Herbicide safeners (*e.g.* naphthalic anhydride, benoxacor, *etc.*) are known to induce the detoxification of herbicides in monocots and were found to increase the P450 content and activate the hydroxylation of herbicides by specific P450s (Zimmerlin and Durst, 1990; Fonne-Pfister and Kreuz, 1990; Persans and Schuler, 1995).

Despite these important roles, however, studies of plant P450s have been impeded by the difficulties in purification of P450s due to their instability and low abundance in plant tissues. Furthermore, a purified P450 protein must be reconstituted with a membrane

electron transport system to prove its enzymatic activity. So far, only a limited number of P450 enzymes and their corresponding cDNAs have been isolated from plants (Song et al., 1993; Chapter I, II, Mizutani et al., 1993; Teutsch et al., 1993; Kraus and Kutchan, 1995; Pan et al., 1995; Koch et al., 1995).

In this study

Plant P450s play important roles in a wide variety of metabolic processes as described above, however, little is known of biochemistry and molecular biology of plant P450s. Purification of a P450 protein and assay of the enzyme activity in a reconstituted system is a straight strategy in order to isolate the corresponding P450 cDNA and investigate physiological significance of the P450. As a first step, I purified a P450 possessing cinnamate 4-hydroxylase activity from mung bean (*Phaseolus aureus*) seedlings by tracing the enzyme activity in a reconstituted system with P450-reductase purified from mung bean, NADPH, and a phospholipid (Chapter I). Second, with the aid of partial amino acid sequences determined for the purified P450, the corresponding cDNAs were isolated from a mung bean hypocotyl cDNA library (Chapter II). To investigate molecular mechanism by which expression of cinnamate 4-hydroxylase gene is regulated in response to environmental stresses, I isolated the cDNA and the gene for the P450 from *Arabidopsis thaliana*, and characterized the expression manner of the P450 gene (Chapter III). Because P450-reductase is an important component of the P450 monooxygenase system, I also investigated the physiological significance of the presence of P450-reductase isoforms in *Arabidopsis thaliana* (Chapter IV). Finally, I demonstrated diversity of P450 genes in a single plant species of *Arabidopsis thaliana* by using molecular biological approaches (Chapter V).

Chapter I

Purification and Characterization of a Cytochrome P450 (Cinnamate 4-Hydroxylase) from Etiolated Mung Bean Seedlings

Introduction

Cinnamate 4-hydroxylase (C4H) catalyzes the hydroxylation of the para position in the aromatic ring of *trans*-cinnamic acid to produce *p*-coumaric acid (Fig. 1, Russel 1971). The hydroxylation of cinnamic acid has long been identified as catalyzed by a P450 and C4H is one of the most abundant P450s in plant tissues.

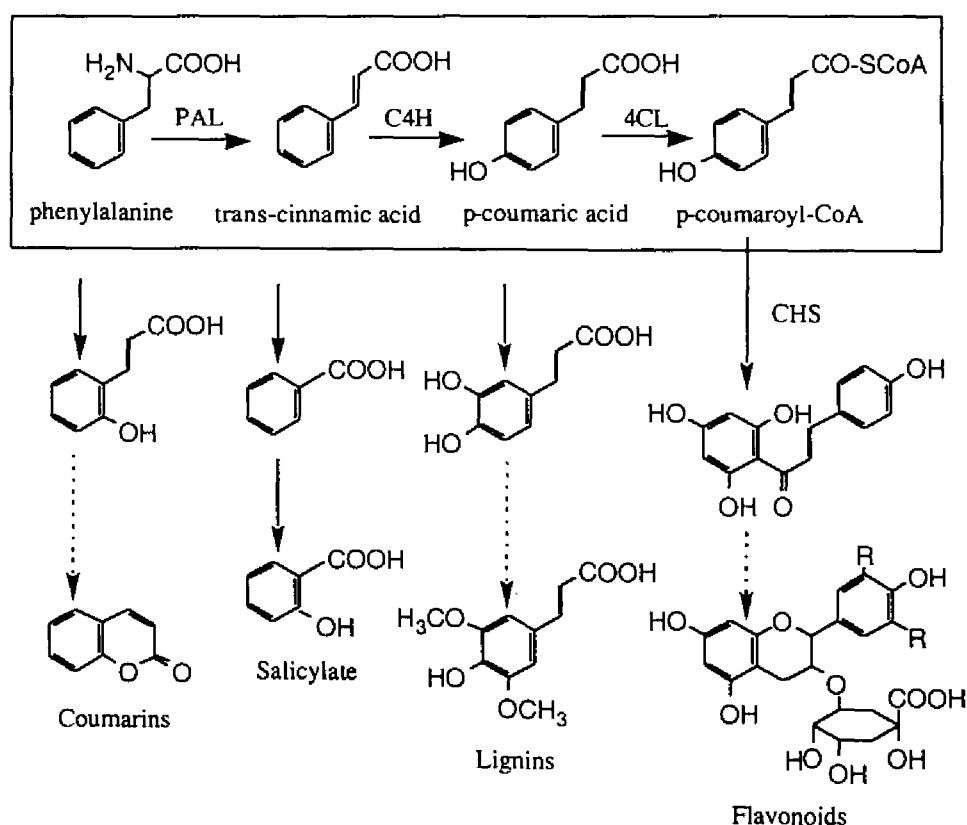


Figure 1. The core reactions of the general phenylpropanoid pathway and related metabolites. The core reactions are boxed.

Together with phenylalanine-ammonialyase (PAL) and 4-coumaroyl:CoA ligase (4CL), it is involved in the core reactions of general phenylpropanoid metabolism which provides phenylalanine-derived intermediates for the biosynthesis of low-molecular-weight flower

pigments, phytoalexins, UV protectants, insect repellents, signal molecules, suberin, lignin, and cell wall components (Hahlbrock and Scheel, 1989). Thus, PAL, C4H, and 4CL are induced in response to changes in environmental conditions or biological damage caused by infection of microorganisms and/or wounding. Phenylpropanoid metabolism is also involved in many aspects of plant cell development and the levels of these three enzymes vary during plant development. However, molecular approaches to studies on the core reactions of phenylpropanoid metabolism and on the inducibility of the enzymes have so far been confined to PAL and 4CL (Hahlbrock and Scheel, 1989).

C4H was first purified from manganese-induced Jerusalem artichoke (*Helianthus tuberosus* L.) (Gabriac et al, 1991). Two P450s possessing 3,9-dihydroxypterocarpan 6a-hydroxylase and C4H activities were also isolated from an elicitor-challenged soybean (*Glycine max*) cell culture (Kochs et al, 1992). In most cases, however, purified C4H showed only insignificant catalytic activities in a reconstituted system with NADPH-P450 reductase.

In order to investigate biochemistry and molecular biology of P450 involvement in phenylpropanoid metabolism, I purified a P450 with C4H activity from microsomes of etiolated mung bean seedlings to electrophoretic homogeneity. The purified P450_{C4H} showed properties characteristic to a P450 and exhibited a high C4H activity in a reconstituted system containing NADPH-P450 reductase purified from mung bean microsomes. Thus, this is the first direct evidence based on a reconstitution study that a physiological C4H activity is associated with a P450 protein.

Materials and Methods

Materials

Sepharose 4B, DEAE Sepharose Fast Flow, 2'5'-ADP Sepharose 4B, Mono S and Mono Q were purchased from Pharmacia, Uppsala. Hydroxyapatite was from Bio-Rad, Richmond. A prepacked hydroxyapatite column (HCA-column, A-5010G) was from Mitsui Toatsu Chemical Inc., Tokyo. An octadecylsilica gel column (TSKgel ODS-120A) was from TOSOH, Tokyo. ω -Aminooctyl Sepharose 4B was prepared according to the method described by Cuatrecasas (1970). Ampure DT columns were obtained from Amersham, Buckinghamshire. Leupeptin, pepstatin, PMSF, and *trans*-cinnamic acid were obtained from Sigma Chemical Co., St. Louis. 1,8-Diaminooctane and Triton X-114 were purchased from Nakarai Chemical Co., Kyoto. Emulgen 913 was a generous gift from Kao-Atlas, Osaka.

Other reagents were of analytical grade. Cytochrome *b5* and NADPH-cytochrome P450 reductase were purified from rabbit liver microsomes according to the methods described by Spatz and Strittmatter (1971) and Yasukochi and Masters (1976), respectively.

Analytical and assay methods

Protein contents was determined according to the method of Lowry et al. (1951) using bovine serum albumin as standard. Since high concentrations of detergents interfered with precise protein determination (Gabriac et al. 1991), samples were treated with Ampure DT columns to remove the detergents and then subjected to TCA precipitation before the protein assay (Cabib and Polacheck 1984). Polyacrylamide slab gel electrophoresis (8-18 % linear gradient) in the presence of 2% SDS was performed by the method of Laemmli (1970). Cytochrome P450 was estimated from CO difference spectrum according to the method of Omura and Sato (1964). *b*-Type cytochromes were estimated from absorbance at α -band in their reduced-minus-oxidized difference spectra using an approximate absorption coefficient of $20 \text{ mM}^{-1}\text{cm}^{-1}$ (Rich and Bendall 1975). All spectrophotometric determinations of cytochromes were carried out at 30 °C in a Shimadzu UV-3101 spectrophotometer.

NADPH-cytochrome P450 reductase was assayed by measuring its NADPH cytochrome *c* reductase activity as described (Imai 1976). C4H activity in microsomal fractions was determined at 30 °C in a reaction mixture (final volume, 0.5 ml) containing 50 mM potassium phosphate (pH 7.25), 200 μM *trans*-cinnamic acid, 1 mM NADPH, 1 mM 2-mercaptoethanol, and microsomal proteins (0.05 mg). The reaction was started by the addition of NADPH. For measurement of C4H activity in a reconstitution system enzyme solution was treated with an Ampure DT column to eliminate detergents and concentrated using Centricon 30 (Amicon) to give P450 concentrations high enough for the assay system. The reconstituted system consisted of 5-10 nM P450, 0.1-0.2 unit/ml of NADPH-P450 reductase, 0.01-0.02% (w/v) sodium cholate, 5-10 $\mu\text{g/ml}$ of dilauroylphosphatidylcholine (DLPC) (sonicated in advance), 50 mM potassium phosphate buffer (pH 7.25), varying concentration of *trans*-cinnamic acid and 0.1 mM NADPH. Practically, concentrated solutions for P450 and NADPH-P450 reductase, sodium cholate and DLPC were mixed first (the volume of this mixture should be less than 50 μl) and then the other components were added in the order given above. In the reconstituted system, concentrations of detergents derived from the enzyme solutions were below their critical micellar concentrations. The reaction was initiated by the addition of

NADPH and run for 10-20 min at 30°C in a total volume of 0.5 ml. The reaction was terminated by the addition of both 20 µl of 2 M HCl and 30 µl of 30% (w/v) TCA. After centrifugation at $10000 \times g$ for 15 min, 0.5 ml of the resultant supernatant was extracted twice with ethyl acetate and the extracts were evaporated to dryness in vacuo. The residue was redissolved in 40 µl of methanol and analyzed by HPLC (Beckman system Gold) equipped with a column of octadecylsilica gel (TSKgel ODS 120A, 4.6×150 mm). The reaction product, *p*-coumaric acid, was separated from the substrate, *trans*-cinnamic acid, by isocratic elution with 50% aqueous methanol containing 0.1% acetic acid. The quantification of *p*-coumaric acid formation was done by comparing the peak area in HPLC analysis with that of authentic *p*-coumaric acid.

Throughout the current purification procedure, C4H was monitored by measuring its C4H activity in the reconstituted system and P450 content was determined from the reduced-CO difference spectrum. Since Jerusalem artichoke C4H showed the type I spectral change upon the addition of *trans*-cinnamic acid (Gabriac et al, 1991), the substrate induced difference spectral shift was employed to identify C4H when necessary.

Preparation of microsomes

Seven-day-old etiolated mung bean seedlings (*Phaseolus aureus*) grown at 20 °C were purchased from a local grocer and then stored at 20°C for 15 h in a dark cabinet maintained at 25°C with relative humidity of 80% to allow induction of C4H. Preliminary experiments indicated that C4H activity was localized in the microsomal fraction of mung bean seedlings and that the level of the activity increased 7-fold with 15 h of the dark treatment. On the other hand, no increase in the C4H activity level was observed when the tissues were stored at 4 °C in the dark. This C4H induction could be reproduced in any batches of seedlings. After this treatment, all experimental procedures were carried out at 4 °C unless otherwise stated.

Whole tissues (400 g) of the seedlings, in which C4H had been induced, were homogenized under N₂ gas stream in 400 ml of 0.2 M potassium phosphate buffer (pH 7.25) containing 0.3 M sucrose, 2 mM EDTA, 2 mM DTT, 0.4 mM PMSF, 2 µM pepstatin, 1 µM leupeptin, 2 mM *trans*-cinnamic acid, 20 mM 2-mercaptoethanol and 10% (w/w) insoluble polyvinylpyrrolidone using a kitchen mixer. The homogenate was squeezed through four layers of cheesecloth and centrifugated at $15,000 \times g$ for 15 min. Microsomes were precipitated from the $15,000 \times g$ supernatant by centrifugation at $100,000 \times g$ for 40 min and

stored at -80 °C in 0.1 M potassium phosphate (pH 7.25) containing 20% (v/v) glycerol, 1 mM EDTA, 1 mM DTT, 0.2 mM PMSF, and 10 mM 2-mercaptoethanol.

Purification of C4H

The stored microsomes (1450 mg protein) were suspended in 700 ml of 0.1 M potassium phosphate buffer (pH 7.25) containing 30% glycerol, 1 mM EDTA, 1 mM DTT and 2 mM 2-mercaptoethanol (buffer A) and centrifugated at $100,000 \times g$ for 60 min. The microsomal pellet was then solubilized using Triton X-114. As described by Bordier (1981), by this solubilization technique, amphiphilic membrane proteins are partitioned predominantly into the detergent-rich phase, whereas hydrophilic molecules are excluded into the detergent-poor or aqueous phase. This method was introduced into the purification of Jerusalem artichoke C4H by Gabriac et al (1991). Practically, microsomes were resuspended in buffer A supplemented with 2% (w/v) Triton X-114 to give a protein concentration of 2 mg/ml and the suspension was stirred for 15 min. The detergent phase was separated from the aqueous phase by centrifugation at $100,000 \times g$ for 60 min. After this solubilization step, C4H was traced throughout the purification process by measuring its activity in the reconstituted system.

The detergent phase thus obtained was diluted 6 times with 20 mM Tris-HCl (pH 8.25) containing 20% (v/v) glycerol, 1 mM DTT, 0.5% (w/v) Emulgen 913 (buffer B) and applied to a column of DEAE-Sepharose Fast Flow (5×40 cm) equilibrated with buffer B. C4H thereby passed through the column, whereas NADPH-P450 reductase, Cytochrome *b5* and Cytochrome *b562* remained bound to the column and could be eluted with a linear KCl gradient (0-500 mM) in buffer B.

The pass-through fraction from DEAE-Sepharose column containing C4H was concentrated 10-fold using Amicon ultrafiltration membrane (YM-30) and dialyzed against 20 mM potassium phosphate (pH 7.25) containing 20% (v/v) glycerol, 1 mM DTT, and 0.5% (w/v) Emulgen 913 (buffer C) for 12 h at 4 °C changing the medium three times 4-h intervals. The dialyzed samples were applied to a hydroxyapatite column (2.6×20 cm) equilibrated with buffer C. The column was extensively washed with a buffer (buffer D) containing the same ingredients as buffer C except that 0.5% (w/v) Emulgen 913 was replaced by 1% (w/v) CHAPS. When A_{278} due to Emulgen 913 disappeared from the column eluate, proteins were eluted with a linear gradient of KCl concentration (0-500 mM) in buffer D. Active fractions

from the hydroxyapatite column were pooled, dialyzed against 50 mM potassium phosphate (pH 7.25) containing 20% (w/v) glycerol, 1 mM DTT and 1% (w/v) CHAPS (buffer E) and applied to an ω -aminooctyl Sepharose 4B column (2.6 \times 15 cm) equilibrated with buffer E. Almost all C4H was recovered in the fraction excluded from the column. The pass-through fraction was concentrated on an ultrafiltration membrane (Amicon, YM30) and diluted with a solution containing 20% (w/v) glycerol, 1 mM DTT, and 0.5% (w/v) Emulgen to give a potassium phosphate concentration of 20 mM. The diluted solution was applied to a packed hydroxyapatite column (HCA-column, A-5010G) equilibrated with buffer C on a linearly increasing potassium phosphate concentration (20-500 mM, pH 7.25) in buffer C at a flow rate of 1 ml/min. Pooled active fractions from the second hydroxyapatite chromatography were diluted 10 times with a solution containing 20% (w/v) glycerol, 1 mM DTT, and 0.5% (w/v) Emulgen 913 and applied to a Mono S column (HR5/5, Pharmacia) equilibrated with buffer C. The column was extensively washed with buffer C and proteins were eluted with a linear gradient of KCl concentration (0-500 mM) in 20 mM Bicine-KOH (pH 8.75) containing 20% (w/v) glycerol, 1 mM DTT, and 0.5% (w/v) Emulgen 913 (buffer F) at a flow rate of 0.5 ml/min. Active fractions from Mono S column chromatography were pooled, dialyzed against buffer F and applied to a Mono Q column (HR5/5, Pharmacia) equilibrated with buffer F. After extensive washing of the column with buffer F from which Emulgen 913 was omitted, C4H was eluted with linear gradient of KCl (0-500 mM) in the same buffer at a flow rate of 0.5 ml/min.

Purification of NADPH-P450 reductase

NADPH-P450 reductase was purified from mung bean seedlings microsomes basically according to the method described by Yasukochi and Masters (1976). Briefly, pooled active fractions from the DEAE-Sepharose column chromatography were dialyzed against 20 mM potassium phosphate (pH 7.25) containing 20% (v/v) glycerol, 0.5% (w/v) Emulgen 913, 0.1 mM DTT, and 1 mM EDTA (buffer G) and applied to a hydroxyapatite column (2.5 \times 5) equilibrated with buffer G. NADPH-P450 reductase was thereby passed through the column. The pass-through fraction was then applied to a 2'5'-ADP Sepharose column (1 \times 7 cm) equilibrated with buffer G. The reductase was eluted from the column with 10 mM potassium phosphate buffer (pH 7.7) containing 20% (v/v) glycerol, 0.2% Emulgen, 1 mM EDTA, 0.1 mM DTT and 0.5 mM NADP. The final preparation was gel electrophoretically homogeneous and had a specific NADPH-P450 reductase activity of 71 unit/mg protein.

Results and discussion

Preparation and solubilization of microsomes

By the dark treatment of mung bean seedlings described in Materials and Methods, the specific C4H activity reached 0.58 nmol *p*-coumarate min⁻¹ mg⁻¹ microsomal protein (Table I). This activity level was almost 7 times of that of the seedlings stored at 4 °C (data not shown), despite the fact that no specific agents such as ethylene were employed to induce the C4H activity.

The mechanism by which the dark treatment causes significant induction of C4H is unclear. The seedlings had been obviously wounded during transportation and displayed at a market at least below 10 °C under illumination. It is possible that the C4H induction in such tissues might be triggered by the increased in the storage temperature to 25 °C. Furthermore, I cannot rule out the possibility that the tissues might have been infected with microorganisms. Fungal infection could induce C4H in peanut and soybean cell cultures (Kochs and Grisebach, 1989; Steffen et al, 1989)

Table I Purification of cinnamate 4-hydroxylase from etiolated mung bean seedlings

Purification step	Cinnamate 4-hydroxylase activity			Cytochrome P450		
	Protein (mg)	Specific activity (nmol min ⁻¹ mg ⁻¹ protein)	Recovery (%)	Total P450 (nmol)	Specific content (nmol/mg protein)	Recovery (%)
Microsomes	1450	0.578	100	140	0.0965	100
Solubilized microsomes	950	n.d.	n.d.	120	0.126	85.7
DEAE Sepharose	445	1.62	86.0	65.0	0.146	46.4
Hydroxyapatite	9.6	74.9	85.7	22.1	2.30	15.7
ω-Aminooctyl Sepharose	6.85	86.3	70.5	18.7	2.72	13.3
Hydroxyapatite	3.23	172	66.2	14.2	4.39	10.1
Mono S	2.08	291	72.1	10.8	5.19	7.71
Mono Q	0.119	861	12.1	1.50	12.6	1.07

The cinnamate 4-hydroxylase activity was determined in the reconstituted system using *trans*-cinnamic acid as a substrate. The assay conditions are described under Materials and Methods. P450 contents were estimated from the CO-difference spectrum according to the method of Omura and Sato (1964). Specific C4H activities are expressed as nmol of *p*-coumarate formed min⁻¹ mg⁻¹ protein.

n.d.: not determined because of the high detergent concentration

Microsomes from the dark-treated seedlings were solubilized with Triton X-114 and the mixture was separated into the detergent and detergent-poor phases. Integral membrane

proteins including C4H were partitioned into the detergent phase as described by Werck-Reichhart et al (1991). Spectral measurement showed that 86% of microsomal P450s was recovered in the detergent-phase without significant formation of P420, an inactive form P450 (Table I). Gabriac et al (1991) reported that C4H activity was almost completely lost when Jerusalem artichoke microsomes were solubilized with Triton X-114 or high concentrations of other non-ionic detergents. I also could not reconstitute C4H activity using our solubilized microsomes (Table I). However, this was not due to inactivation of C4H, but to a high detergent concentration that interfered with the reconstitution of C4H activity. In fact, C4H activity could be measured in the reconstituted system throughout the following chromatography steps in which detergents other than Triton X-114 were employed at lower concentrations. In addition, careful removal of detergents from samples containing P450 was essential to measure C4H activity in the reconstituted system.

When Emulgen 913, Triton X-100, CHAPS, and sodium cholate were used for solubilization at any protein/detergent ratios, no satisfactory results were obtained in the subsequent purification steps, owing to inactivation of C4H or incomplete solubilization.

Purification of C4H

A typical purification of mung bean C4H is summarized in Table I. By six steps of column chromatography, C4H was purified 1490-fold from the microsomal fraction with an overall yield of 12.1%. The specific P450 content in the final preparation was 12.6 nmol/mg protein. It gave a single protein band with an apparent molecular mass of 58-kDa on SDS-PAGE (Fig. 2). This molecular mass of mung bean C4H is close to those reported for C4Hs from other plant tissues (Gabriac et al, 1991; Kochs et al, 1992).

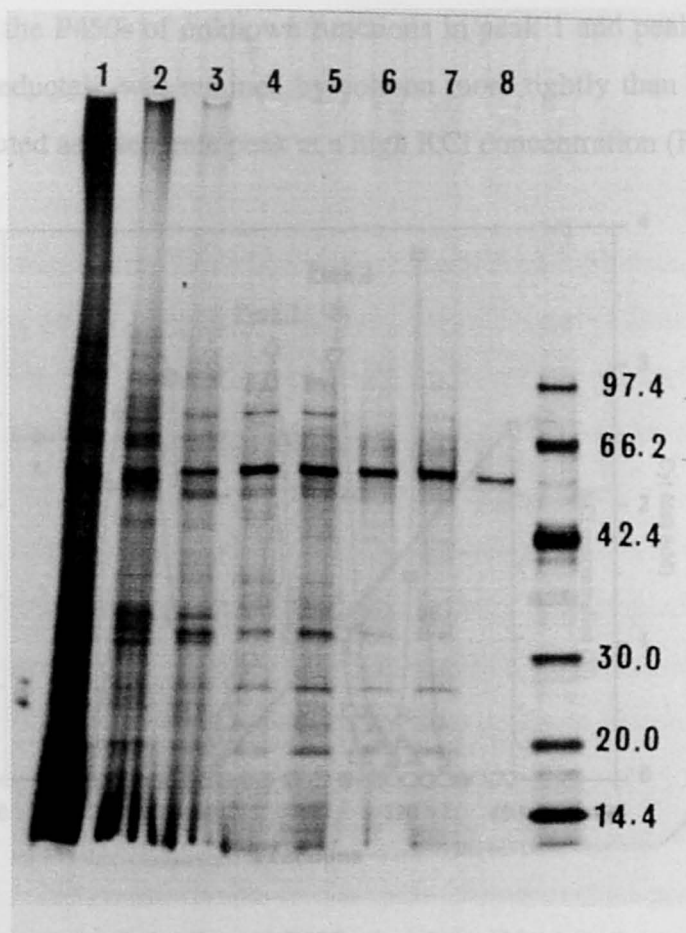


Fig. 2 SDS-PAGE of fractions from the chromatographic steps of the purification of P450C4H. Gel electrophoresis was performed according to the method of Laemmli (1970) using 8-18% polyacrylamide slab gel. The proteins were visualized by silver staining. Lane 1, Microsomes (4 μ g protein); lane 2, the detergent phase obtained from solubilized mung bean microsomes (2 μ g protein); lane 3, DEAE-Sepharose passed-through fraction (1 μ g protein); lane 4, the first hydroxyapatite elute (0.3 μ g protein); lane 5, -aminooctyl-Sepharose pass-through fraction (0.3 μ g protein); lane 6, the second hydroxyapatite elute (0.1 μ g protein); lane 7, Mono S elute (0.1 μ g protein); lane 8, Mono Q elute (0.03 μ g protein); the right lane, molecular weight markers.

DEAE Sepharose chromatography of the diluted solubilized microsomes, the first chromatographic step of the purification procedure, resulted in the recovery of 86% of C4H activity and 46% of P450 in the pass-through fraction (Table I, Fig. 3), leaving about 40% of P450 still retained by the column. The retained P450 was eluted from the column with a KCl gradient in two prominent peaks (Fig. 3, peak 1 and 2). In these fractions, however, neither C4H activity, nor the type I spectral change induced by 100 μ M *trans*-cinnamic acid was observed (data not shown), indicating that the microsomes contain at least two other forms of P450 having no C4H activity. Cytochrome *b*562 and *b*5 were eluted from the column

concomitantly with the P450s of unknown functions in peak 1 and peak 2, respectively (Fig. 3). NADPH-P450 reductase was retained by column more tightly than the P450s and *b*-type cytochromes and eluted as a separate peak at a high KCl concentration (Fig. 3).

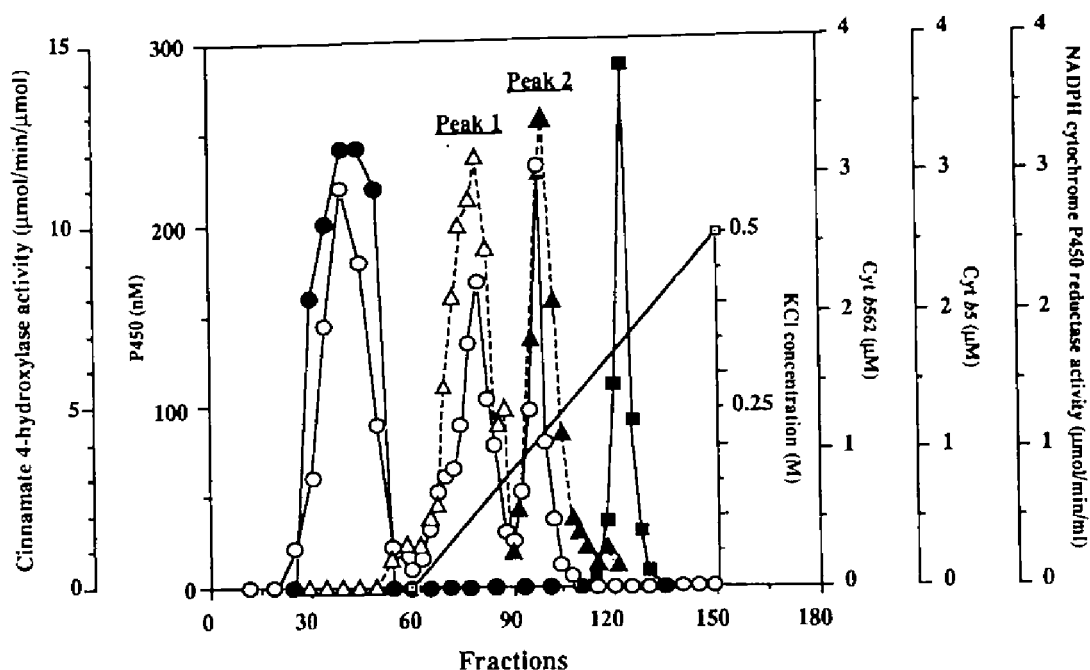


Fig. 3 Elution patterns of cytochrome P450, cinnamate 4-hydroxylase, cytochrome *b*562, cytochrome *b*5, and NADPH-P450 reductase during DEAE-Sepharose Fast Flow column chromatography.

The solubilized detergent-phase obtained from mung bean microsomes (1.45 g protein) was diluted and applied to the column (5 x 40 cm) and 20 ml of fractions were collected. Proteins were eluted with a KCl gradient (0-500 mM) in a buffer consisting of 20 mM Tris-HCl (pH 8.25), 20% glycerol, 1 mM DTT, and 0.5% Emulgen 913. P450 concentration was determined from the CO difference spectrum according to the method of Omura and Sato (1964). Cytochrome *b*562 and *b*5 were estimated from the reduced-minus-oxidized absorption spectra according to the method of Rich and Bendall (1975).

●, cinnamate 4-hydroxylase; ○, P450; △, cytochrome *b*562;
▲, cytochrome *b*5; ■, NADPH-P450 reductase.

The first hydroxyapatite column chromatography, which resulted in almost 50-fold purification of C4H, was the most effective step in the current purification procedure (Table I). ω -Aminoethyl Sepharose column chromatography was an essential step in the purification procedure, although significant purification of C4H was not achieved (Table I, Fig. 2). If this step was omitted, no clear purification of C4H was possible during the following chromatographic steps, probably because of hydrophobic aggregation of C4H with other membrane proteins. Interactions with the hydrophobic ω -aminoethyl group may have prevented the aggregate formation leading to successful purification of C4H. Three more

chromatographic steps were necessary after ω -aminooctyl Sepharose chromatography to purify C4H to homogeneity.

Reconstitution of C4H activity

C4H activity was reconstituted by mixing the purified C4H with mung bean NADPH-P450 reductase in the presence of DLPC and cholate. The reconstituted activity ($861 \text{ nmol min}^{-1} \text{ mg}^{-1} \text{ C4H protein} = 38 \text{ nmol min}^{-1} \text{ nmol}^{-1} \text{ P450}$, Table II) was the highest level of C4H activity so far reported. Namely, C4H purified from elicitor challenged soybean cell cultures showed a reconstituted activity of $5139 \text{ nkat (kg/protein)}$ which corresponds to $25 \text{ pmol min}^{-1} \text{ nmol}^{-1} \text{ P450}$ (Kochs et al, 1992). On the other hand, the activity could not be reconstituted with C4H purified from Jerusalem artichoke by taking advantage of the type I spectral change induced by the addition of *trans*-cinnamic acid (Gabriac et al, 1991). Thus, our result is the first clear evidence from the reconstituted system proving that a physiological activity is intrinsically associated with a higher plant P450 protein (P450_{C4H}).

Table II The reconstitution of cinnamate 4-hydroxylase activity

Assay condition	Cinnamate 4-hydroxylase activity ($\text{nmol min}^{-1} \text{ nmol}^{-1} \text{ P450}$)
Complete I (mung bean NADPH-P450 reductase)	38.8
II (rabbit NADPH-P450 reductase)	38.4
+ 0.1 nmol rabbit Cytochrome <i>b5</i> ^a	37.9
- NADPH ^b	n.d.
- NADPH-P450 reductase ^b	n.d.
- C4H ^b	n.d.

Values are means of duplicated experiments. The complete reconstituted system contained 5 nM purified P450_{C4H}, 0.1 unit ml^{-1} NADPH-P450 reductase from mung bean or rabbit, 0.01% (w/v) sodium cholate, $10 \mu\text{g ml}^{-1}$ dilauroylphosphatidylcholine, 0.1 mM NADPH, and 0.4 mM *trans*-cinnamic acid. The specific P450 content of P450_{C4H} used for activity determination was $6.8 \text{ nmol mg}^{-1} \text{ protein}$.

^a Rabbit liver cytochrome *b5* was added to complete assay mixture at a concentration of 0.1 μM .

^b NADPH, NADPH-P450 reductase, or P450_{C4H} was omitted from the complete assay mixture, respectively.

The reconstituted system specifically produced *p*-coumaric acid and no other metabolites were detected by HPLC (not shown). The C4H activity reconstituted with rabbit liver NADPH-P450 reductase was virtually identical with that obtained with the mung bean reductase (Table II). Thus, mung bean C4H can receive electrons from rabbit liver NADPH-P450 reductase as efficiently as from mung bean reductase, suggesting that the reductase's structures responsible for the interaction with P450s are conserved between higher plants and mammals. The addition of purified rabbit cytochrome *b5* did not affect the C4H activity (Table II). It is, however, not certain whether mung bean C4H activity is not activated by cytochrome *b5* or rabbit liver cytochrome *b5* cannot interact with the mung bean microsomal electron transport system. The K_m value for *trans*-cinnamic acid was determined to be 1.8 μ M. The value is very similar to those reported for Jerusalem artichoke and soybean C4H activities determined in microsomes (Gabriac et al, 1991; Kochs et al, 1992).

Spectral properties of P450C4H

The absolute absorption spectra of P450C4H are shown in Fig. 4.

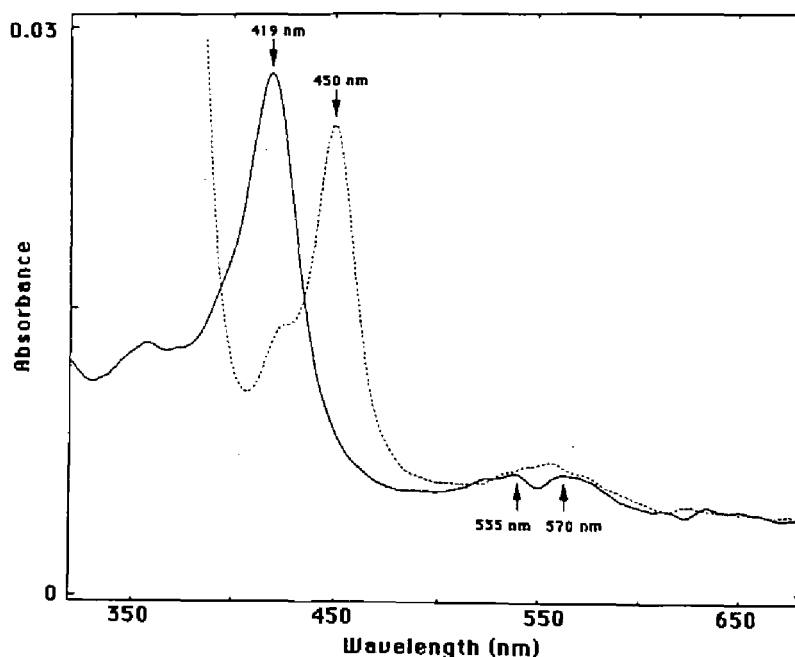


Figure 4 Absorption spectra of the purified P450C4H. The purified preparation was dissolved (0.1 μ M) in 50 mM potassium phosphate buffer (pH 7.25) containing 20% glycerol and 0.5% Emulgen 913. To reduce P450, a few grains of sodium dithionite was added to the same solution.

—————, oxidized form; , reduced CO-complex.

In the oxidized state, the P450 exhibited a low-spin type spectrum having a Soret absorption maximum at around 419 nm and α - and β -bands at 570 nm and 535 nm, respectively. the reduced CO-complex of C4H showed a Soret absorption maximum at 450 nm. When *trans*-cinnamic acid was added to the oxidized form of C4H, the spectrum was partially converted to the high spin state as reported for Jerusalem artichoke C4H (Gabriac et al, 1991), indicating the binding of *trans*-cinnamic acid to the substrate binding site of this P450.

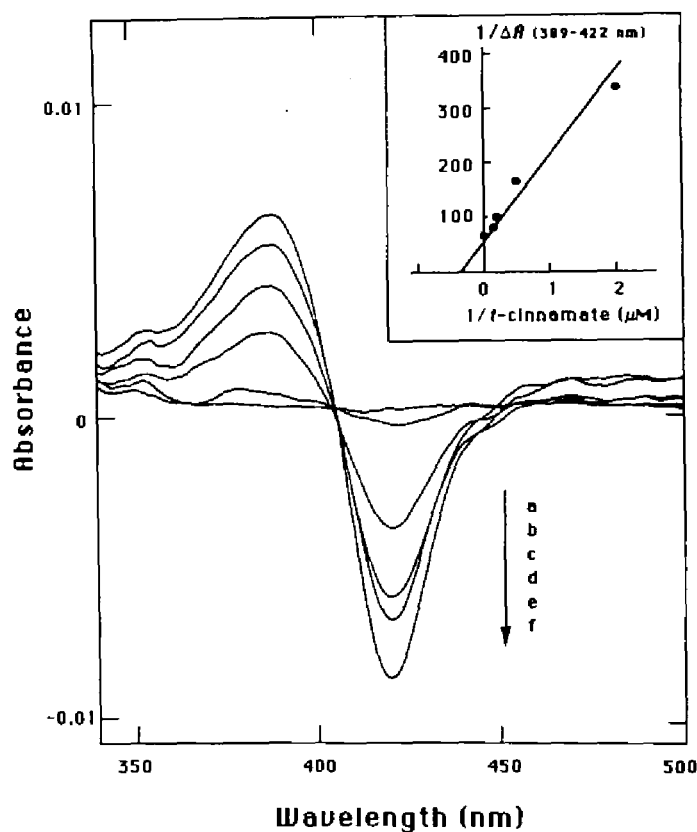


Figure 5 *trans*-Cinnamic acid induced difference spectra of P450C4H. The purified preparation (0.2 μ M) was dissolved in 50 mM potassium phosphate buffer (pH 7.25) containing 20% glycerol and 0.5% Emulgen 913. The final concentrations of *trans*-cinnamic acid were (a), 0; (b), 0.5; (c), 2; (d), 5; (e), 10, and; (f), 50 μ M. Spectral dissociation constant was calculated from a double reciprocal plot of absorbance A (389-422 nm) versus the substrate concentration.

Fig. 5 shows type I substrate binding difference spectra induced by the addition of *trans*-cinnamic acid in a concentration range from 0 to 50 μ M. The dissociation constant of *trans*-cinnamic acid for C4H calculated from spectral titration was 2.8 μ M, and was almost identical to the value reported for Jerusalem artichoke C4H. The spectral dissociation constant is consistent with the K_m value (1.8 μ M) for *trans*-cinnamic acid obtained in the reconstitution

system, indicating that the cinnamate binding to the oxidized C4H is the first step of the overall cinnamate 4-hydroxylation.

Summary

A cytochrome P450 (P450_{C4H}) possessing *trans*-cinnamate 4-hydroxylase (C4H) activity was purified to apparent homogeneity from microsomes of etiolated mung bean seedlings. Upon SDS-polyacrylamide gel electrophoresis, the purified preparation gave a single protein band with a molecular mass of 58-kDa. Its specific P450 content was 12.6 nmol mg⁻¹ protein. Using NADPH as an electron donor, the purified C4H aerobically converted *trans*-cinnamic acid to *para*-coumaric acid with a specific activity of 38 nmol min⁻¹nmol⁻¹ P450 in a reconstituted system containing NADPH-cytochrome P450 reductase purified from the seedlings or rabbit liver microsomes, dilauroylphosphatidylcholine, and NADPH. This specific activity is by far the highest for reconstituted C4H system so far reported and provides direct evidence that C4H activity is actually associated with a P450 protein. In the oxidized state C4H showed a typical low-spin type absorption spectrum with a Soret peak at 419 nm. A partial spectral shift to the high spin state was observed when *trans*-cinnamic acid was added to the oxidized C4H. By spectral titration, the dissociation constant of the cinnamic acid-P450 complex was determined to be 2.8 μ M. This value is similar to the K_m value (1.8 μ M) for *trans*-cinnamic acid determined in the reconstituted system.

Chapter II

Molecular Cloning and Sequencing of a cDNA Encoding Mung Bean Cytochrome P450 (P450C4H) Possessing Cinnamate 4-Hydroxylase Activity

Introduction

C4H, together with phenylalanine ammonia-lyase (PAL) and 4-coumaroyl:CoA ligase (4CL), is involved in the core reactions of general phenylpropanoid metabolism, providing precursors for the biosynthesis of secondary metabolites such as lignins and suberin (Hahlbrock and Scheel, 1989). However, compared to PAL and 4CL, much less is known about P450C4H gene expression. As described in Chapter I, I have purified C4H from mung bean seedlings and demonstrated the reconstitution of C4H activity with the purified protein. Here I report the isolation and sequencing of cDNA clones for P450C4H from a mung bean hypocotyl cDNA library with the aid of partial amino acid sequences determined for the purified C4H.

Materials and Methods

Plant materials

P450C4H was purified as described in Chapter I (Mizutani et al. 1993) from etiolated mung bean (*Phaseolus aureus*) seedlings purchased from a local grocer in Japan. For the isolation of total RNA, mung bean seedlings were germinated on damp vermiculite in dark at 25 °C. Before the isolation of total RNA, P450C4H expression was induced by incubating 3-day-old etiolated hypocotyl sections for 24 hr at 25 °C in an airtight vessel containing 25 mM potassium phosphate (pH 6.0) supplemented with 2% (w/v) sucrose and 0.005% (w/v) chloramphenicol.

Amino acid sequencing

After digestion of purified P450C4H with lysyl endopeptidase, peptide fragments were separated on Mono Q column (Pharmacia). Three peptides were chosen and purified by

reverse phase HPLC using a μ PRC C2/C18 column on SMART system (Pharmacia). Amino acid sequences were analysed by automated Edman degradation using a Simadzu protein sequencer (PSQ-2) with an on-line Simadzu amino acid analyzer (PTH-1) equipped with a Wako-Pack WS-PTH column (4.6×250 mm, Wako Pure Chemical). The N-terminal sequence was determined directly from the purified P450C_{4H}.

Construction of cDNA Library

Total RNA was prepared from mung bean hypocotyls by phenol/chloroform extraction followed by lithium chloride precipitation as described (Lagrimini et al, 1987). Poly(A)⁺ RNA was isolated from total RNA using a poly(A)⁺ Quick mRNA isolation kit (Stratagene). A cDNA library was constructed from the poly(A)⁺ mRNA using a ZAP cDNA synthesis kit (Stratagene).

Polymerase Chain Reaction and Isolation of cDNA Clones

First strand cDNA was synthesized from 1 μ g of mung bean hypocotyl poly(A)⁺ RNA and a part of this was used as a template for polymerase chain reaction (PCR), which was carried out in 100 μ l of 10 mM of Tris-HCl (pH 8.3) containing degenerate primers (each at 10 μ M), 200 μ M dATP, 200 μ M dCTP, 200 μ M dTTP, 200 μ M dGTP, 1.5 mM MgCl₂, 50 mM KCl, 25 units/ml of Amplitaq DNA polymerase (Perkin Elmer/Cetus). The reaction was performed through 35 cycles of 30 sec. at 94 °C, 30 sec at 46 °C and 90 sec at 72 °C using a Perkin Elmer/Cetus thermal cycler model 480. PCR products were separated by agarose gel (2%) electrophoresis. A major band (1.3-kb) was isolated from the gel and cloned into the pCRII vector using a TA cloning kit (Invitrogen). The 1.3-kb fragment was labeled with [³²P]dCTP by random priming method (Feinberg and Vogelstein, 1983), and about 120,000 plaques from the cDNA library were screened with the labeled fragment as a probe.

Other Methods

The inserts of positive cDNA clones were excised in vivo as subclones in the pBluescript plasmid (Stratagene) and sequenced as double strand templates using the dideoxynucleotide chain termination method (Hattori et al, 1986). Northern hybridization analysis of hypocotyl poly(A)⁺ RNA was performed as described (Payne et al, 1990)) using the insert of one of the positive cDNA clones as a probe. Sequence comparison was made using the programs GAP and PILEUP (Devereaux et al, 1984)

Results and Discussions

C4H, purified from etiolated mung bean seedlings, showed significant cinnamate 4-hydroxylase activity (38 nmol *p*-coumarate formed min⁻¹ nmol⁻¹ P450) with NADPH as an electron donor in a reconstitution system containing mung bean NADPH-P450 reductase and dilauroylphosphatidylcholine (Mizutani et al, 1993). We determined the amino acid sequences of the C4H N-terminal portion (49 residues) as well as three peptides from a lysyl endopeptidase digest (Fig. 2, underlined).

Based on the sequences thus determined, several degenerate oligonucleotide primers for PCR were synthesized. PCR was performed using, as a template, first strand cDNA synthesized from mung bean hypocotyl poly(A)⁺ RNA. Since the location of the three peptides sequenced in C4H polypeptide was unknown, PCR was carried out using two primers, one derived from N-terminal portion and the other from an internal peptide segment. When MA22 (corresponding to residues 41-46 in the N-terminal portion) and a mixture of MA23 and MA24 (corresponding to peptide γ) were used as primers (Fig. 1A), a 1.3-kb fragment was obtained as a major product (Fig. 1B).

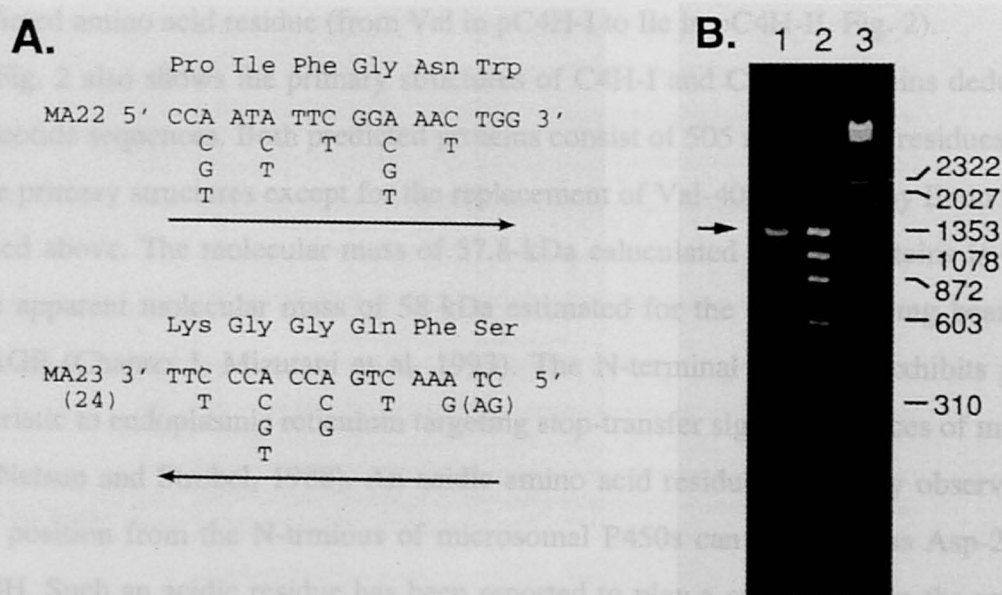


Figure 1. PCR amplification of a fragment of P450C4H cDNA. (A) The DNA sequences of degenerate primers MA22, MA23, and MA24, aligned with the peptide sequences from which they were designed. (B) Ethidium bromide-stained agarose gel showing the products of PCR amplification from first strand cDNA. An arrow indicates the 1.3-kb fragment that was subcloned for sequence analysis. Size markers are given in kb.

The primary structure deduced from the nucleotide sequence of the 1.3-kb fragment contained one of the internal sequences determined above (D-Y-F-V-D-E-R-K, peptide β in Fig. 2) and a sequence (F-G/S-X-G-X-R/H-X-C-X-G/A/D, double underlined in Fig. 2) that has been reported to be highly conserved in all the P450s so far sequenced (HR2 or heme-binding region) (Gotoh and Fujii-Kuriyama, 1989). These findings indicated that this fragment actually reflects a portion of C4H sequence. Therefore this fragment was labeled and used as a probe for screening a mung bean cDNA library.

The two of eight strongly hybridized clones (pC4H-I and pC4H-II) contained cDNA inserts significantly long to represent the full-length C4H message, and were completely sequenced. As shown in Fig, the pC4H-I sequence consisted of a 1515-bp open reading frame, a 25-bp 5'-untranslated region, a 223-bp 3'-noncoding region, and a poly(A) tail. A poly(A) addition signal (AATAAA) is seen 20-bp upstream of the poly(A) stretch. The nucleotide sequence of pC4H-II was slightly different from that of pC4H-I. Thus, pC4H-II sequence lacks one A at the 5'-terminus and six A's in the poly(A) tail compared to pC4H-I. Furthermore, both G at nucleotide 143 (coding region) and T at nucleotide 1766 (noncoding region) are replaced by A in the pC4H-II sequence. The substitution at 143 leads to change in the predicted amino acid residue (from Val in pC4H-I to Ile in pC4H-II, Fig. 2).

Fig. 2 also shows the primary structures of C4H-I and C4H-II proteins deduced from the nucleotide sequences. Both predicted proteins consist of 505 amino acid residues and have the same primary structures except for the replacement of Val-40 in C4H-I by Ile in C4H-II as mentioned above. The molecular mass of 57.8-kDa calculated for the proteins is consistent with the apparent molecular mass of 58-kDa estimated for the purified mung bean C4H by SDS-PAGE (Chapter I, Mizutani et al, 1993). The N-terminal structure exhibits properties characteristic to endoplasmic reticulum targeting stop-transfer signal sequences of microsomal P450s (Nelson and Strobel, 1988). An acidic amino acid residue frequently observed at 2nd and 3rd position from the N-terminus of microsomal P450s can be found as Asp-2 in mung bean C4H. Such an acidic residue has been reported to play a crucial role in the retention of microsomal P450s at cytoplasmic surface of endoplasmic reticulum (Belzer et al, 1990). The four partial amino acid sequences determined from the purified C4H (peptide α , β , γ , and δ in Fig. 2) can be found in the deduced primary structures. However, residues Val-38 and Ile-498 determined by peptide sequencing are replaced by Leu and Val, respectively, in the deduced primary structure. We conclude that the two cDNA clones isolated in this study actually encode C4H, although minor differences are seen between the two clones and purified C4H for

unknown reasons. Conceivably the mung bean genome contains a small family of C4H genes, individual members of which could account for the differences in sequence observed. Genomic Southern blot analysis is consistent with this possibility (data not shown).

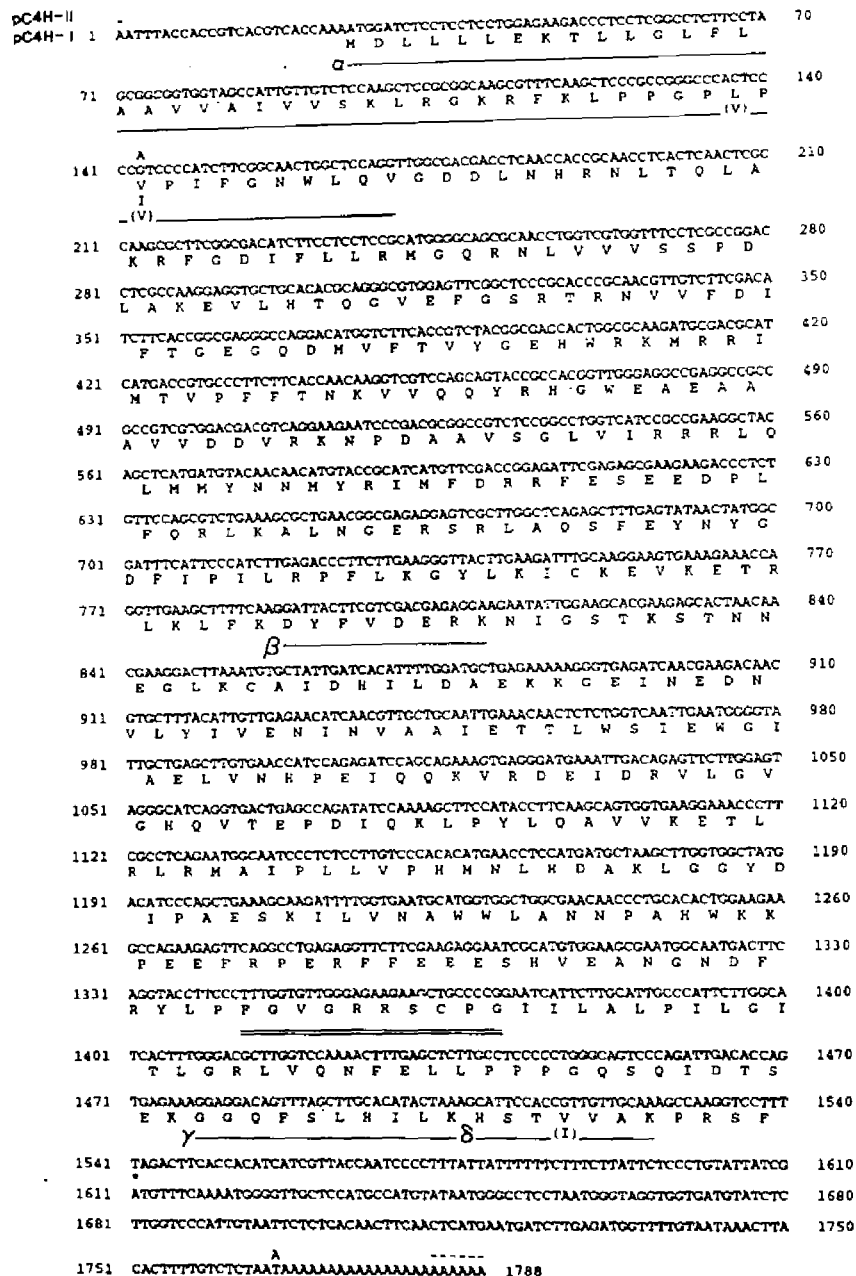


Figure 2. Sequence of P450C4H cDNA clone pC4H-I. Regions of the predicted protein sequence corresponding to the protein chemically determined peptide sequences are underlined. The individual peptide sequences are labeled (α , β , γ , δ). Differences between the predicted protein sequences and peptides α and δ are indicated in parentheses. Hyphens above the sequence indicate deletions in pC4H-II compared to pC4H-I. The two single nucleotide changes in pC4H-II compared to pC4H-I are also shown above the sequence; one of these results in an amino acid change which is shown below the predicted protein sequence. The conserved putative heme-binding region is double underlined. The sequence has been deposited in DDBJ, EMBL, and GenBank as accession L07634.

Summary To determine the size of C4H transcripts, poly(A)⁺ RNA from mung bean hypocotyls was subjected to Northern blot analysis using the labeled insert of pC4H-I as a probe. As shown in Fig. 3, a single mRNA of about 1.8-kb in size was detected, which indicates that the cDNA clones analysed in this study are very close to the full-length of the mRNA.

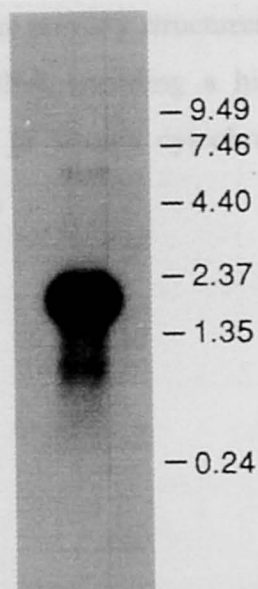


Figure 3. Blot hybridization of mung bean poly(A)⁺ RNA probed with [³²P]-labeled pC4H-I cDNA insert. The positions of RNA size standards are indicated, and their sizes are given in kb.

Three cDNAs for higher plant P450s have so far been cloned and sequenced. These are an avocado P450 involved in fruit ripening (Bozak et al, 1990) and two closely related periwinkle P450s of unknown functions (Vetter et al, 1992). Comparison of the predicted protein sequence of mung bean C4H (pC4H-I sequence) with those predicted for these higher plant P450s showed that C4H sequence is 31.2% and 23.5% identical with those of avocado P450 and periwinkle P450s, respectively. Comparison was also made between mung bean C4H and P450s belonging to the known P450 gene families, but none of them were found to be more than 40% identity to the C4H sequence. These results indicate that mung bean C4H belongs to a novel P450 gene family that is distinct from those so far reported, and the C4H was designated CYP73A2.

Summary

With the aid of partial amino acid sequences determined for cinnamate 4-hydroxylase (P450C_{4H}) purified from mung bean seedlings, two cDNA clones were isolated and their inserts were completely sequenced. The nucleotide sequences of the two clones were nearly identical and contained an open reading frame predicted to encode a polypeptide consisting of 505 amino acid residues. The partial amino acid sequences determined for the purified P450C_{4H} closely corresponded to the primary structures deduced from the cDNA sequences. This is the first isolation of a cDNA encoding a higher plant P450 possessing a clear physiological activity. Comparison to known cytochromes P450 indicated that P450C_{4H} belongs to a novel P450 gene family.

Chapter III

Isolation of a cDNA and a Genomic Clone Encoding Cinnamate 4-Hydroxylase from *Arabidopsis thaliana* and its Expression Manner *in Planta*

Introduction

The genes involved in the phenylpropanoid pathway (Fig. 1) have been isolated and characterized in several plant species (Mol et al., 1988; Hahlbrock and Scheel, 1989; Dixon and Harrison, 1990; Dangl, 1992; Dixon and Paiva, 1995).

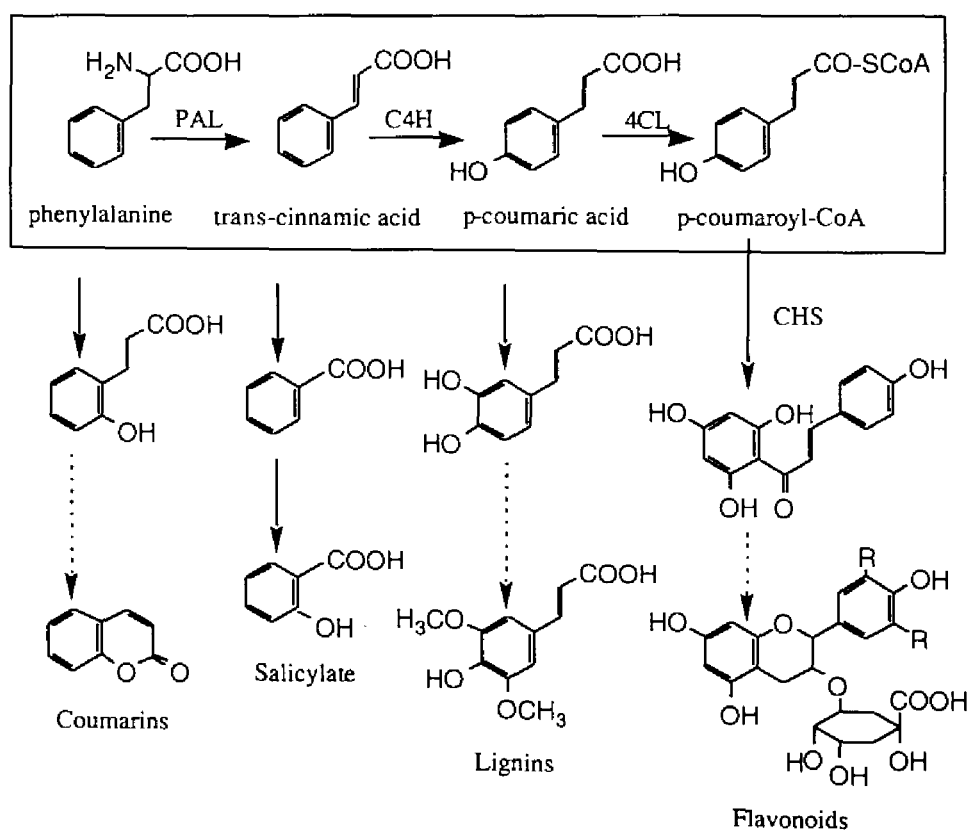


Figure 1. The core reactions of the general phenylpropanoid pathway and related metabolites. PAL, phenylalanine ammonia-lyase; C4H, cinnamate 4-hydroxylase; 4CL, 4-coumarate:CoA ligase; CHS, chalcone synthase. The core reactions are boxed.

Specifically, it has been known that phenylalanine ammonia-lyase (PAL) and 4-coumarate:CoA ligase (4CL) are coordinately expressed during development and in response

to biotic and abiotic stresses such as pathogen infection (Lawton and Lamb, 1987), wounding (Lois and Hahlbrock, 1992), and UV irradiation (Chappell and Hahlbrock, 1984; Douglas et al., 1987). *In vivo* footprinting studies have led to the identification of three sequence motifs (boxes P, A, and L) in the promoter regions of parsley *PAL1* and *4CL* genes for responding to UV irradiation and elicitor treatment (Lois et al., 1989). These sequence motifs have been also found in most of the known *PAL* and *4CL* gene promoters (Logemann et al., 1995). It is likely that the levels of *PAL* and *4CL* are regulated to meet the cellular requirements for phenylpropanoid derived-secondary metabolites (Dixon and Paiva, 1995). However, it is still obscure how *C4H* expression is regulated in concert with other genes of phenylpropanoid pathway.

As described in Chapter I and II, I have purified the *C4H* protein (Mizutani et al., 1993a) and isolated a *C4H* cDNA from mung bean seedlings (Mizutani et al., 1993b). *C4H* cDNAs have been also isolated from artichoke (Teutsch et al., 1993) and alfalfa (Fahrendorf and Dixon, 1993). In this chapter, I describe the isolation of the *C4H* cDNA from *Arabidopsis thaliana* and characterization of its expression patterns in response to wounding and light. RNA gel blot analysis indicated that *C4H* was expressed coordinately with other enzymes of the phenylpropanoid pathway, *PAL1* and *4CL*, implicating a common transcriptional activation mechanism for these genes. Sequence analysis of a genomic clone containing the 5'-promoter region of the *C4H* gene revealed that the cis-elements (boxes P, A, and L) involved in the regulated expression of *PAL* and *4CL* genes (Logemann et al., 1995) also exist in the *C4H* promoter.

Materials and Methods

Plant materials/treatments

Arabidopsis thaliana ecotype Columbia (Col-0) (Lehle Seeds, Tucson, AZ) seedlings were grown under a sterile condition on 0.8% agarose plates containing GM medium (Valvekens et al., 1988) in a growth chamber maintained at 22°C under the continuous light or a light/dark cycle (9/15h).

For wounding treatment, leaves were harvested from 3-week-old plants grown under continuous light, cut into 2 mm in width, and incubated for 1 to 9h under continuous light in a petri dish containing GM medium and 0.005% (w/v) chloramphenicol. For the 0h time

point, sliced leaves were frozen immediately without further incubation under continuous light.

Isolation of cDNA and genomic DNA clones

Total RNA was prepared from 7-d-old seedlings by phenol/chloroform extraction, followed by lithium chloride precipitation as described by Lagrimini et al. (1987), and poly(A)⁺ RNA was isolated from the total RNA using a poly(A)⁺ Quick mRNA isolation kit (Stratagene, La Jolla, CA). A cDNA library was constructed from the poly(A)⁺ RNA in a bacteriophage vector λ ZAPII (Stratagene, La Jolla, CA) using a Uni-ZAP XR Gigapack II Gold cloning kit (Stratagene) according to the manufacturers' instruction.

The cDNA of mung bean C4H (Chapter II, Mizutani et al., 1993b) was labeled with [³²P]-dCTP by the random priming labeling method (Feinberg and Vogelstein, 1983). A total of 600,000 phages from the *Arabidopsis* cDNA library were plated, and duplicate lifts were prepared on nylon filters (Dupont, Boston, MA). The filters were hybridized and washed under low stringency conditions: hybridization, 16h at 45°C in a hybridization buffer containing 1% BSA, 7% SDS, 50 mM sodium phosphate (pH 7.5) and 1 mM EDTA; washing, 10 min in 6 × SSC supplemented with 0.1% SDS at room temperature and 20 min in 2 × SSC with 0.1% SDS at 37°C. Twenty-six positive signals were identified by autoradiography, and the positive plaques were isolated performing an additional round of screening. Bluescript SK⁻ phagemids were isolated by *in vivo* excision with R408 helper phage (Stratagene). Since partial DNA sequencing of the positive clones indicated that the DNA inserts of different lengths were derived from a transcript of a single P450 gene, the DNA sequence of the longest insert (DDBJ accession number: D78596) was completely determined. According to the criteria of the cytochrome P450 nomenclature committee (Nelson et al., 1993), the P450 was designated CYP73A5.

Genomic DNA clones were isolated by using a 315-bp *EcoRI/BamHI* fragment from the N-terminal region of the CYP73A5 cDNA as a probe. A total of 50,000 plaques from a λ EMBL3 SP6/T7 genomic library of *Arabidopsis thaliana* ecotype Columbia (Clontech, Palo Alto, CA) was screened under high stringency conditions: hybridization, 16h at 65°C in the same hybridization buffer as described above; washing, 10 min in 2 × SSC with 0.1% SDS at room temperature and 30 min in 0.1 × SSC containing 0.1% SDS at 65°C. Six positive plaques were isolated through two additional rounds of screening. Restriction mapping

analysis indicated that the cloned inserts were all identical except for the different length of the 5'- or 3'-noncoding region. One of the positive clones contained the longest insert of 6-kb including a part of the *CYP73A5* gene, and a 1.2-kb *Bam*HI/*Xho*I fragment from the insert was subcloned into a Bluescript SK⁻ plasmid. The DNA sequence of the fragment (DDBJ accession number: D78597) was completely determined.

Heterologous expression in insect cells

The CYP73A5 protein was expressed using the baculovirus expression vector system basically according to the method described previously (Summers and Smith, 1987), using a baculovirus transfer vector pVL1392 (Invitrogen, San Diego, CA), *Spodoptera frugiperda* 21 (Sf21) cells (Invitrogen) and an infectious BaculoGold Baculovirus DNA (Pharmingen, San Diego, CA). Sf21 cells were maintained at 27°C as a monolayer culture in a Grace's medium supplemented with 0.33% TC yeastolate, 0.33% lactoalbumin, 10% fetal bovine serum, 50 µg/ml of gentamycin sulfate, 200 µM 5-aminolevulinic acid and 200 µM ferrous citrate. The expressed CYP73A5 protein was partially purified from the infected Sf21 cells. Briefly, the infected cells were sonicated and centrifuged at 100,000 × g for 1h. The pellet was homogenized with buffer A containing 20 mM potassium phosphate (pH 7.25), 20% glycerol and 1 mM dithiothreitol, and proteins were solubilized in buffer A supplemented with 1% Emulgen 913 (Kao Atlas, Tokyo, Japan) (buffer B). After centrifugation at 100,000 × g for 1h, the supernatant was applied to a DEAE-Sepharose column (5 × 10 cm) equilibrated with buffer B. The pass-through fraction from the DEAE-Sepharose column was applied to a hydroxyapatite column (2 × 5 cm) equilibrated with buffer B. The column was washed with buffer A until the absorption at 280 nm derived from Emulgen 913 disappeared from the eluted buffer. After the removal of the detergent, the protein was eluted from the column with 0.3 M potassium phosphate buffer containing 20% glycerol and 1 mM dithiothreitol. P450 was monitored during the chromatography from the absorbance at 420 nm derived from a Soret absorption peak of the oxidized form of the P450. C4H activity was reconstituted as described in Chapter I (Mizutani et al., 1993a). Briefly, 5 nM of the purified recombinant CYP73A5 protein was mixed in a reaction mixture containing 50 mM potassium phosphate buffer (pH 7.2), 0.1 unit/mL NADPH-cytochrome P450 reductase purified from mung bean seedlings (Chapter I, Mizutani et al., 1993a), 0.001% (w/v) sodium cholate, 10 mg/mL dilauroylphosphatidylcholine (Funakoshi, Tokyo, Japan), 0.4 mM *trans*-

cinnamic acid and 0.1 mM NADPH. The reaction was started by adding NADPH. *p*-Coumaric acid formation was determined by HPLC according to the method described in Chapter I (Mizutani et al., 1993a). Cytochrome P450 was estimated from the CO difference spectrum (Omura and Sato, 1964).

DNA preparation and DNA blot analysis

Genomic DNA was isolated from shoots of 3-week-old Arabidopsis seedlings and purified by ethidium bromide-CsCl density gradient centrifugation as described by Ausubel et al. (1987). For Southern blot, genomic DNA was digested with the indicated restriction enzymes, separated by 0.7% agarose gel electrophoresis, and transferred to a nylon membrane. The membrane was hybridized with the full-length CYP73A5 cDNA as a probe and washed under the high stringency conditions as described above.

RNA preparation and RNA blot analysis

Total RNA was isolated by phenol/chloroform extraction followed by lithium chloride precipitation as described by Lagrimini et al. (1987). Samples of total RNA (5 µg) were separated by 1.2% formaldehyde-agarose gel electrophoresis, and blotted onto a nylon membrane. The full-length CYP73A5 cDNA or the gene specific probes described below were used for hybridization. Hybridization and washing were carried out under the high stringency conditions described above. Intensity of hybridization signals was quantified using an imaging analyzer BAS2000 (Fuji Film, Tokyo, Japan). The hybridized DNA probes were stripped from the membranes by boiling in a buffer containing 0.01× SSC and 0.01% SDS for 20 min, and blots were re-hybridized to an Actin-1 probe (Nairn et al., 1988) for normalization.

Specific probes for *PAL1*, *PAL2*, *PAL3* (Wanner et al., 1995), *4CL* (Lee et al., 1995), and *CHS* (Feinbaum and Ausubel, 1988) genes were prepared from Arabidopsis genomic DNA by PCR. 5 sets of primers were designed from the specific regions of the genes:

PAL1, 5'-GATCTTGTAATCTCCTCTTAGTTAATCTT-3'

and 5'-GGATCCTCCGCGTTGATCACCATGTTCTTT-3';

PAL2, 5'-AGATCTGATCTCATTACCTAAACACAAAC-3'

and 5'-GCTAGCTTTACACCGGCTCTTGAAGTCTC-3';

PAL3, 5'-CCGGCTCGTTATATAAAAGCTCCAGACTTG-3'

and 5'-CCCGAGCCTCCTCAGAAAGCTCCACTGTCTG-3';
4CL, 5'-TTACAATGGCGCCACAAGAACAAGCAGTTT-3'
and 5'-GGAGATGTTTTGGAAGATGTAGTCGTGGAG-3';
CHS, 5'-ATACTATAATGGTGATGGCTGGTGCTTCTT-3'
and 5'-CATGCGCTTGAACCTTCTCCTTGAGGTCGGT-3'.

Each of the amplified PCR products was cloned into a pCRII vector using a TA cloning kit (Invitrogen). DNA sequences of the PCR products were determined to be identical to those of the corresponding genes.

DNA sequencing and analysis

DNA sequencing was performed using a DyeDeoxy™ Terminator Cycle Sequencing kit (Applied Biosystems, CA) and an automated DNA sequencer (Applied Biosystems 373A). Sequence analysis was performed using a software, DNASIS Ver. 3.5 (Hitachi Software Engineering America, Ltd., San Bruno, CA). The nucleotide sequence data of *PAL1* (L33677), *PAL2* (L33678), *PAL3* (L33679) (Wanner et al., 1995), and 4CL (U18675) (Lee et al., 1995) of *Arabidopsis thaliana* were retrieved from the EMBL, GenBank, and DDBJ Nucleotide Sequence Databases using the corresponding accession numbers shown in the parentheses.

Results

Isolation and characterization of CYP73A5 cDNA

A cDNA clone was isolated from an *Arabidopsis* cDNA library, using the C4H cDNA from mung bean (Chapter I, Mizutani et al., 1993b) as a probe. The insert of the clone consists of a 1515-bp open reading frame, a 56-bp 5'-untranslated region, a 140 bp 3'-noncoding region and a poly(A) tail (not shown). The primary structure deduced from the cDNA sequence shows general properties of microsomal P450s from different organisms including higher plants (Fig. 2). Thus it contains a 30-amino acid long hydrophobic stretch at N-terminus, which is a signal-anchor sequence to retain P450s on the cytoplasmic surface of the endoplasmic reticulum membrane (Nelson and Strobel, 1988). The N-terminal signal-anchor sequence is followed by a proline rich region, which is thought to be involved as an α -helix breaker in the process of correct folding of microsomal P450s and is also important

in heme incorporation into the P450s (Yamazaki et al., 1993). Furthermore, the predicted protein contained a sequence (F-G/S-X-G-X-R/H-X-C-X-G/A/D, underlined in Fig. 2) that has been known to be highly conserved in all the P450s so far sequenced (HR2 or heme-binding region) (Gotoh and Fujii-Kuriyama, 1989).

Arabidopsis	1	MDLLLLLEKSL	IAVFVAVILA	TVISKLRGKK	LKLPPGPIPI	PIFGNWLQVG
Mung bean	1	-----T-	LGL-L-AV-	I-V-----R	F-----L-V	-----
Artichoke	1	-----I--T-	V-L-A-I-G-	IL-----	F-----V	-----
Alfalfa	1	-----T-	L-L-I-ATI-	VT-----R	F-----V	-----Y----
Arabidopsis	51	DDLNHRNLVD	YAKKFGDLFL	LRMGQRNLVV	VSSPDLTKEV	LLTQGVVEFGS
Mung bean	51	-----TQ	L--R--i--	-----	-----A--	-H-----
Artichoke	51	-----T-	L--R--EIL-	-----	-----E-A--	-H-----
Alfalfa	51	-----T-	---R--EI--	-----	-----E-A--	-H--C-----
Arabidopsis	101	RTRNVVFDIF	TGKGQDMVFT	VYGEHWRKMR	RIMTVPFPTN	KVVQQNREGW
Mung bean	101	-----	--E-----	-----	-----	-----Y-H--
Artichoke	101	-----	-----	-----	-----	-----Y-Y--
Alfalfa	101	-----	-----	-----	-----	-----Y-Y--
Arabidopsis	151	EFEAASVVED	VKKNPDSATK	GIVLRKRLQL	MMYTNMFRIM	FDRRFESEDD
Mung bean	151	-A---A--D-	-R---A-VS	-L-I-R---	---N--Y---	-----E-
Artichoke	151	-A---A--D-	-----AA--E	---I-R---	---N--Y---	-----E-
Alfalfa	151	-S--E---N-	--N-AEASVG	---I-----	---NI-Y---	-----E-
Arabidopsis	201	PLFLRLKALN	GERSRLAQSF	EYNYGDFIPI	LRPFLRGYLK	ICQDVKDRRI
Mung bean	201	---Q-----	-----	-----	-----K---	--EK--ET-L
Artichoke	201	---K-----	-----	-----	-----N---	--EK--K--
Alfalfa	201	---VK-----	-----	-----	-----K---	--EK--L
Arabidopsis	251	ALFKKYFVDE	RKQIASSKPT	GS.EGLKCAI	DHILEAEQKG	EINEDNVLYI
Mung bean	251	K---D-----	--N-G-T-S-	.NN-----	---D--K--	-----
Artichoke	251	Q---D-----	--K-G-T-KM	DNNQ.-----	---KE--	-----
Alfalfa	251	Q---D-----	--KLE-T-S-	T-ND-----	---D-QK--	---D-----
Arabidopsis	300	VENINVAATIE	TTLWSIEWGI	AELVNHPEIQ	SKLGNELDTV	LGPGVQVTEP
Mung bean	300	-----	-----	-----	Q-VRD-I-R-	--V-H-----
Artichoke	300	-----	-----	-----	A--RH-----K	-----I---
Alfalfa	301	-----	-----	-----QD--	N-VRE-M-R-	-----H-----
Arabidopsis	350	DLHKLPYLQA	VVKETLRLRM	AIPLLVPHMN	LHDAKLAGYD	IPAESKILVN
Mung bean	350	-IQ-----	-----	-----	-----G-	-----
Artichoke	350	-VQN-----	-----	-----	-----G-F-	-----
Alfalfa	351	-----	-I-----	-----	---P--N-F-	-----
Arabidopsis	400	AWWLANNPNS	WKKPEEFRPE	RFFEEESHVE	ANGNDFRYVP	<u>FGVGRRSCPG</u>
Mung bean	400	-----AH	-----	-----	-----L-	-----
Artichoke	400	-----DQ	-----	--L--AK--	-----L-	-----
Alfalfa	401	-----P--AH	-----	--L-----	-----L-	-----
Arabidopsis	450	IILALPILGI	TIGRMVQNFE	LLPPPGQSKV	DTSEKGGQFS	LHILNHSIIV
Mung bean	450	-----	-L--L-----	-----QI	-----	---K--TV-
Artichoke	450	-----	---L-----	-----I	--D-----	---K--T--
Alfalfa	451	-----	---L--V-	-----I	-----	---K--T--
Arabidopsis	500	MKPRNC				
Mung bean	500	A---SF				
Artichoke	500	A---SF				
Alfalfa	501	A---SF				

Figure 2. Deduced amino acid sequence of the CYP73A5 protein from Arabidopsis with those of the C4Hs from mung bean, artichoke, and alfalfa.

The predicted amino acid sequence of CYP73A5 is aligned to those of C4H proteins from mung bean, artichoke, and alfalfa. Identical amino acid residues are indicated by dashes, and dots are inserted to maximize the sequence homology. Heme-binding domain (HR2 region) is underlined.

The amino acid sequence of the P450 protein was 84.7%, 85.5%, and 83.2% identical to those of the C4Hs from mung bean (Mizutani et al., 1993b), artichoke (Teusch et al., 1993), and alfalfa (Fahrendorf and Dixon, 1993), respectively (Fig. 2). No P450s subfamilies other than *CYP73* exhibited identity greater than 40%. According to the criteria of the cytochrome P450 nomenclature committee (Nelson et al., 1993), the clone was designated *CYP73A5*.

Heterologous expression of *CYP73A5* in insect cells

Although the deduced protein structure of *CYP73A5* is 84.7% identical to that of mung bean C4H (Fig. 2), a single amino acid substitution could change substrate specificity of P450s belonging to the same gene subfamily (Lindberg and Negishi, 1989; Johnson, 1992). In other words, it is possible that *CYP73A5* could encode a P450 with an activity different from the para-hydroxylation of *trans*-cinnamate. Therefore, expression of the *CYP73A5* protein in insect cells using the baculovirus expression system and characterization of the *CYP73A5* protein was necessary.

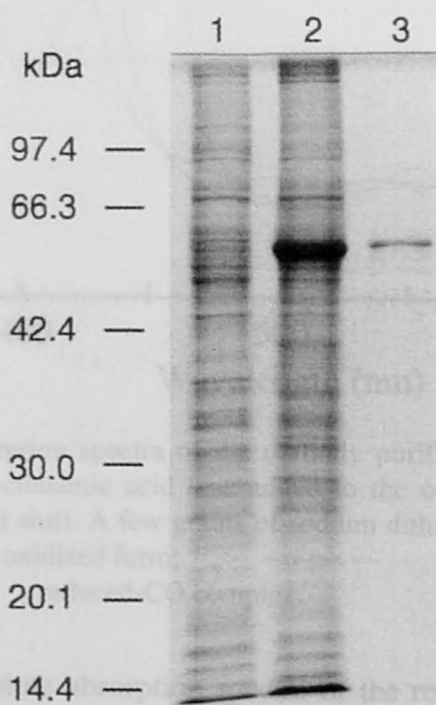


Figure 3. Expression of recombinant *CYP73A5* protein.

SDS-polyacrylamide gel electrophoresis was performed using 12% polyacrylamide slab gel, and proteins were visualized with Coomassie Brilliant Blue R-250.

Lane 1, microsomes of mock-infected Sf21 cells (10 mg protein); lane 2, microsomes of Sf21 cells (10 mg protein) infected with the recombinant virus containing *CYP73A5* cDNA; lane 3, partially purified *CYP73A5* (10 pmoles P450) through a two-step column chromatography using DEAE-Sepharose and hydroxylapatite, which yielded 40 nmol of the *CYP73A5* from the microsomal fraction containing 160 nmol of total P450s. Cytochrome P450 was determined from the reduced CO-difference spectrum.

On SDS-PAGE analysis, a new intense band of 58 kDa appeared in the microsomal fraction of the recombinant virus infected-cells (Fig. 3, lane 2). This protein band was not found in that of the mock-infected cells (Fig. 3, lane 1). The molecular mass of 58 kDa of this protein (57.7 kDa) is consistent with that calculated from the CYP73A5 primary structure.

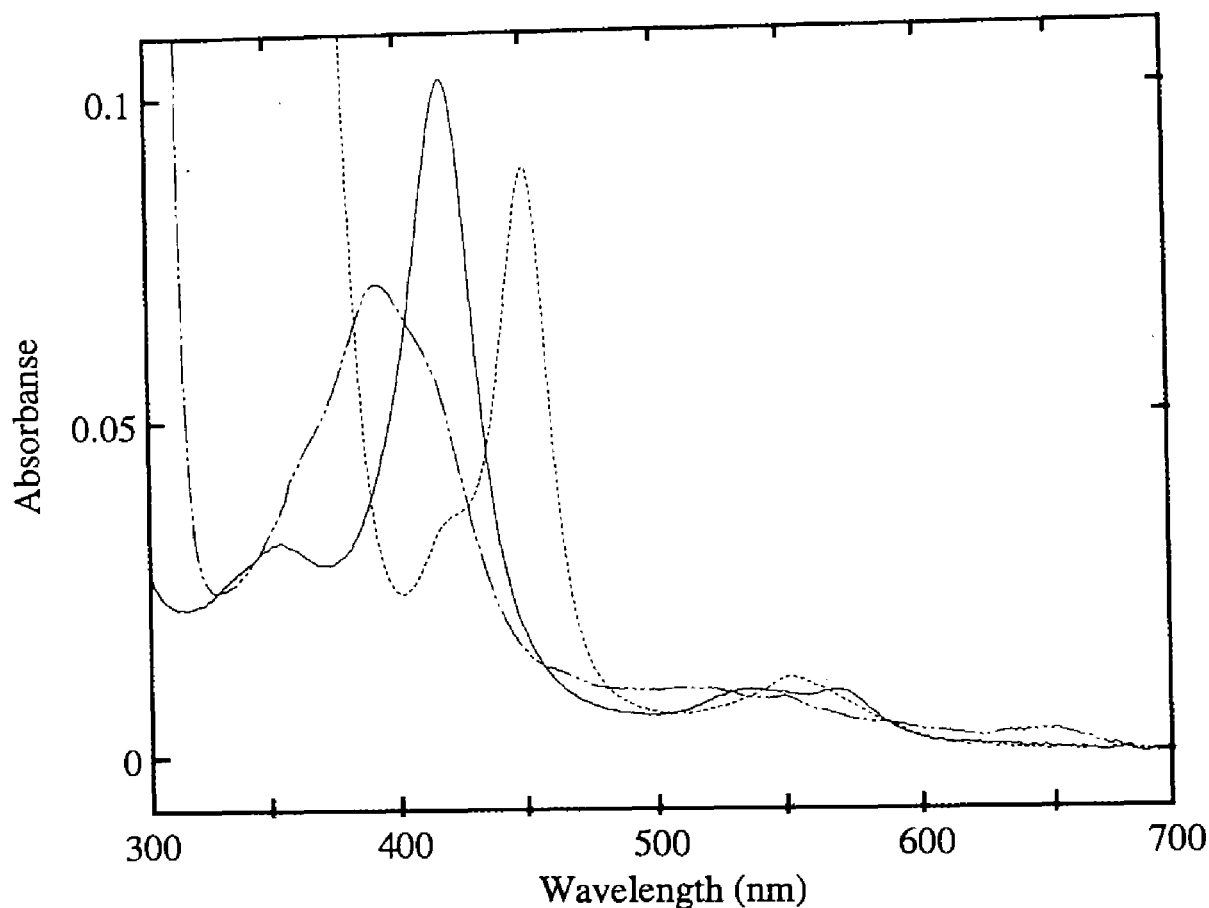


Figure 4. Absolute absorption spectra of the partially purified recombinant CYP73A5 protein. 1 mM of *trans*-cinnamic acid was added to the oxidized form to obtain the substrate induced-spectral shift. A few grains of sodium dithionite were added to reduce the P450. —, oxidized form; ---, + *trans*-cinnamic acid; , reduced-CO complex.

Fig. 4 shows the absolute absorption spectra of the recombinant CYP73A5 protein. The oxidized form exhibited a Soret absorption peak at around 420 nm, and the reduced CO-complex showed a Soret absorption maximum at 450 nm. The addition of a putative substrate, *trans*-cinnamic acid, to the oxidized form caused the shift of the absorption peak from 420 to 390 nm. This result indicated that *trans*-cinnamic acid bound to the CYP73A5 protein as the substrate, thereby modifying the heme electronic state from the low-spin to the high-spin (Ruckpaul et al., 1989). Other potential substrates for plant P450s such as laurate,

nerol, geraniol, and ferulate did not cause such a spectral shift observed with *trans*-cinnamic acid.

C4H activity was reconstituted with the CYP73A5 protein, NADPH-P450 reductase from mung bean seedlings, dilauroylphosphatidylcholine, and *trans*-cinnamic acid as described in Chapter I (Mizutani et al., 1993a). In the reconstituted system, the CYP73A5 protein catalyzed the hydroxylation of *trans*-cinnamic acid to produce *p*-coumaric acid. Formation of *p*-coumaric acid was dependent on both the amount of the CYP73A5 protein and the reaction time. Any other reaction products were not detected under the experimental condition. The reconstituted activity (68 nmol/min/nmol P450) was higher than that of the C4H protein purified from mung bean seedlings (38 nmol/min/nmol P450) (Chapter I, Mizutani et al., 1993a). These results clearly indicated that the CYP73A5 gene encodes the cinnamate 4-hydroxylase (C4H) of Arabidopsis.

Genomic Southern blot analysis

Southern blot analysis was performed to estimate the number of the CYP73A5 genes in Arabidopsis genome (Fig. 5).

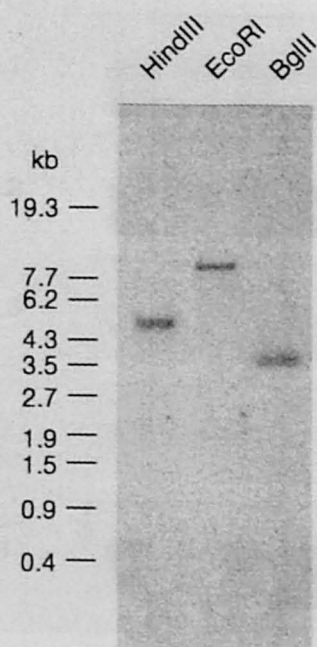


Figure 5. Southern blot analysis of the CYP73A5 gene.

Arabidopsis thaliana Columbia genomic DNA (5 µg) was digested with the indicated restriction enzymes, separated on 0.7% agarose gel, blotted on a nylon membrane, and hybridized with a [³²P]-labeled full-length CYP73A5 cDNA. The migration of size makers is shown to the left of the blot.

Genomic DNA was digested with three restriction enzymes, *Bgl*III, *Eco*RI, and *Hind*III, which do not cut the insert of the CYP73A5 cDNA, and the full-length CYP73A5 cDNA was used as a probe. A single hybridization signal was observed in each digestion, and the identical hybridization pattern was obtained under the low stringency condition (data not shown). These results indicated that *CYP73A5* exists as a single copy gene in the Arabidopsis genome.

Expression patterns in Arabidopsis

Since it was shown that the *CYP73A5* actually encoded the C4H of Arabidopsis, we compared its expression pattern with those of the genes involved in the phenylpropanoid pathway (Fig. 1) by RNA gel blot analysis (Figs. 6, 7, and 8).

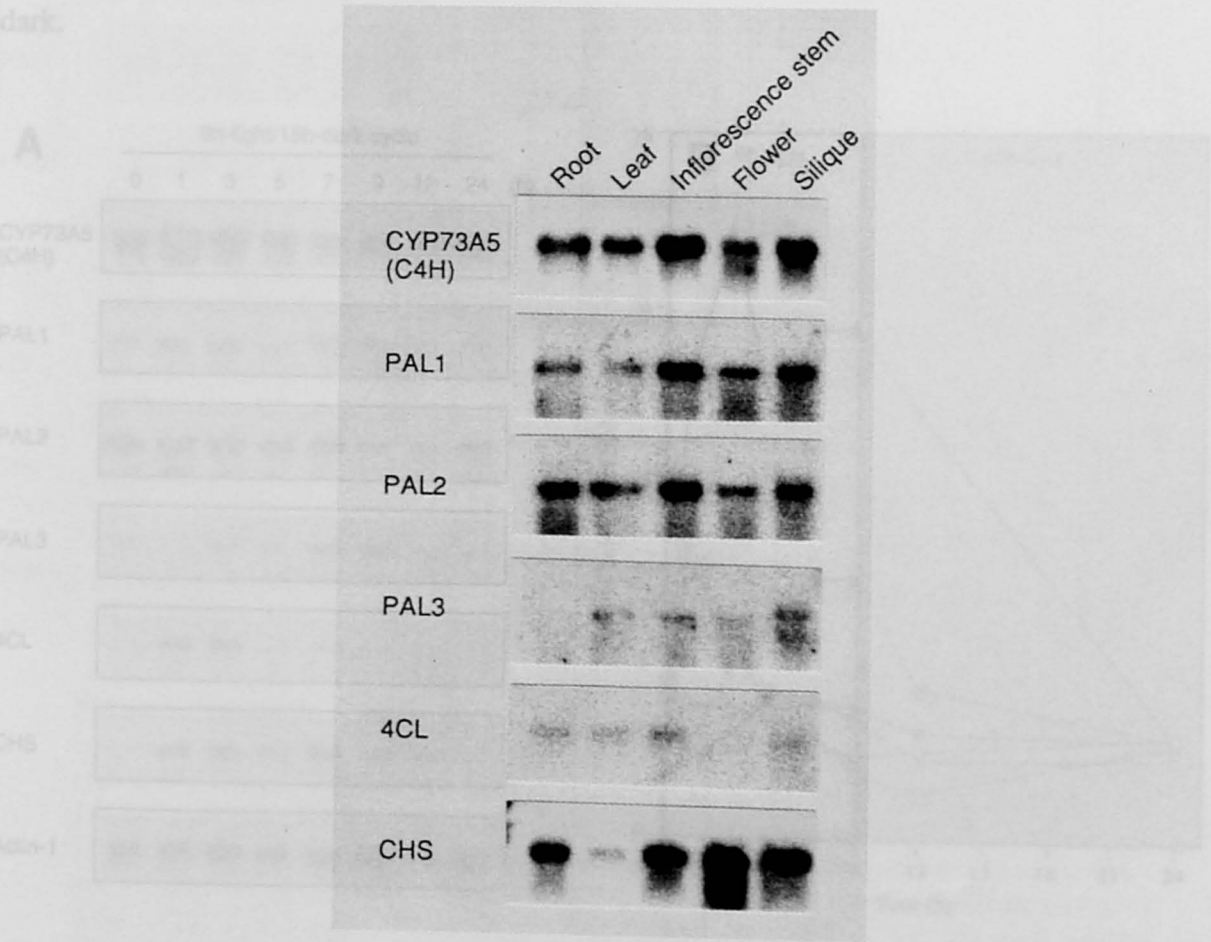


Figure 6. Tissue specific expression of Arabidopsis *PAL1*, *PAL2*, *PAL3*, *CYP73A5* (*C4H*), *4CL*, and *CHS* genes.

Total RNA was isolated from the roots and leaves of 3-week-old plants, from the inflorescence stems and flowers of 4-week-old plants, and from the siliques of 5-week-old plants. Plants were grown under the continuous light. Five μ g total RNA was separated on formaldehyde gels, transferred to nylon membranes, and hybridized to the indicated probes.

The *CYP73A5* (*C4H*) expression level was highest in inflorescence stems and was significantly higher in roots and siliques than in leaves and flowers. Expression patterns of *PAL1*, *PAL2*, and *4CL* were essentially the same as that of *CYP73A5* (*C4H*). In contrast, *PAL3* transcript could not be detected in roots, and *CHS* was most strongly expressed in flowers.

Expression of these genes were also studied under the light/dark cycle (9/15h) (Fig. 7). The levels of *CYP73A5* (*C4H*), *PAL1*, and *4CL* transcripts transiently increased about 3-fold within 1h of the onset of the light period, and decreased continually to a basal level. The amount of *PAL2* mRNA was slightly increased in the light. *PAL3* transcript level gradually increased and reached a maximum level in 9h under the light and rapidly decreased within 3h in the dark. *CHS* transcription was induced 6-fold in the light and gradually decreased in the dark.

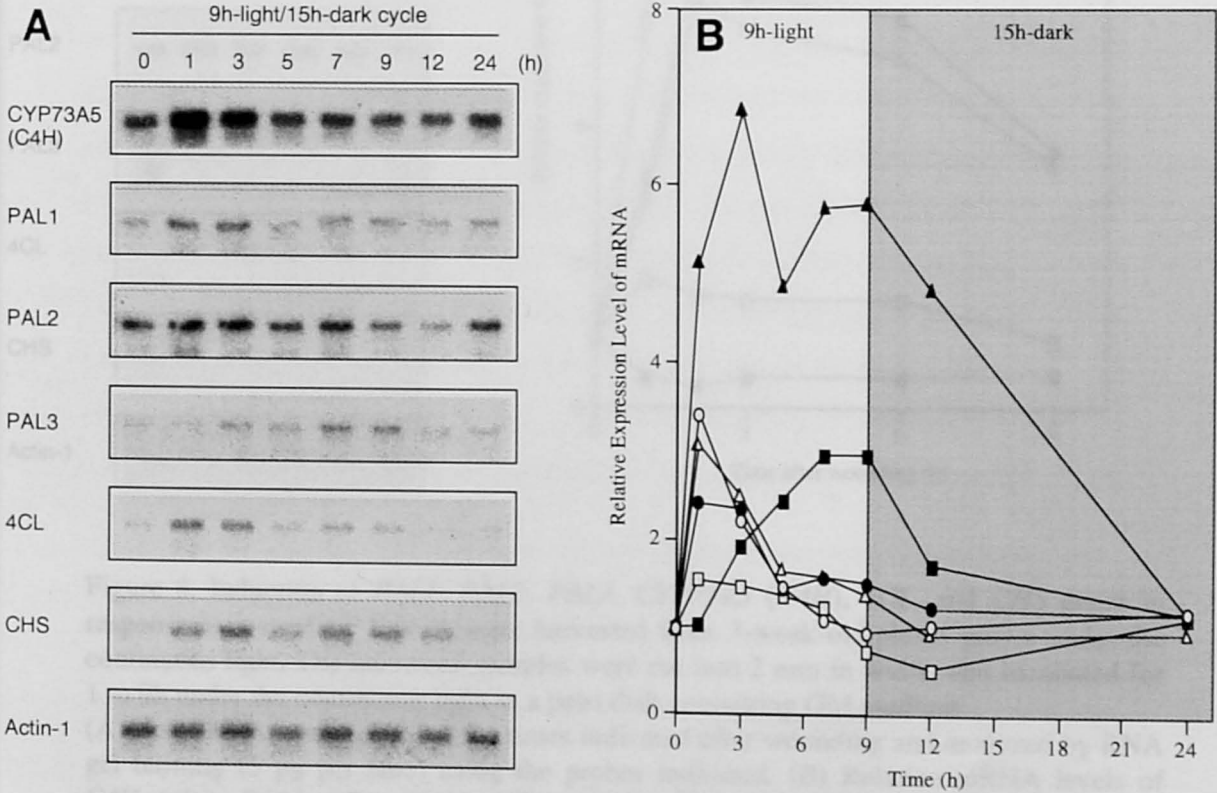


Figure 7. Changes in expression levels of *PAL1*, *PAL2*, *PAL3*, *CYP73A5* (*C4H*), *4CL*, and *CHS* genes under a 9-h light/15-h dark cycle: light period, 0-9h; dark period, 9-24h. (A) Total RNA was isolated from the leaves of 3-week-old plants at times indicated after the onset of light period (0h), and 5 mg of the samples were analyzed by RNA gel blots using the probes indicated. (B) Relative mRNA levels of *C4H* (○), *PAL1* (●), *PAL2* (□), *PAL3* (■), *4CL* (△), and *CHS* (▲). The hybridization signals were quantified using an imaging analyzer and were normalized relative to *Actin-1* (Nairn et al., 1988) expression.

Fig. 8 shows effects of wounding on the expression of these genes. *CYP73A5* (*C4H*), *PAL1*, and *4CL* were significantly induced within 1h of wounding. The expression levels of these genes reached maxima within 2h of the treatment and then decreased. *PAL2* mRNA level was also slightly increased, while *PAL3* and *CHS* mRNA levels were decreased by the wounding treatment.

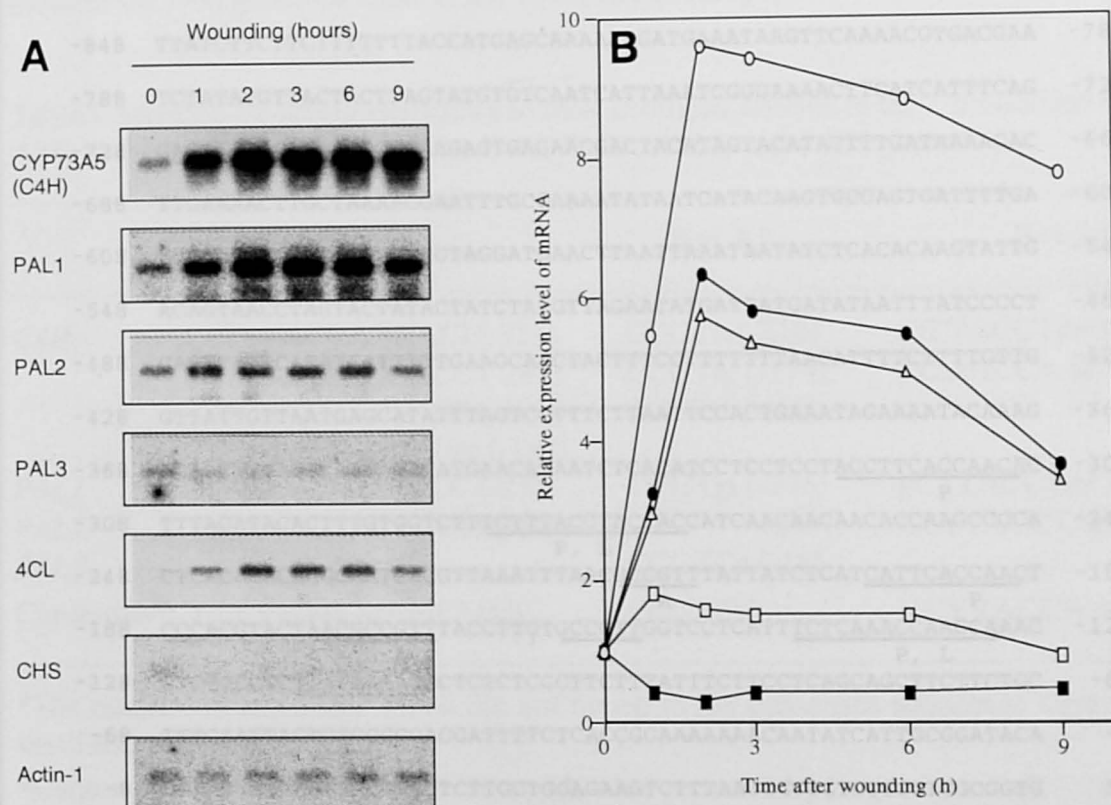


Figure 8. Induction of *PAL1*, *PAL2*, *PAL3*, *CYP73A5* (*C4H*), *4CL*, and *CHS* genes in response to wounding. Leaves were harvested from 3-week-old plants grown under the continuous light. The harvested samples were cut into 2 mm in width, and incubated for 1 to 9h under the continuous light in a petri dish containing GM medium

(A) Total RNA was isolated at the times indicated after wounding and analyzed by RNA gel blotting (5 μ g per lane) using the probes indicated. (B) Relative mRNA levels of *C4H* (○), *PAL1* (●), *PAL2* (□), *PAL3* (■), and *4CL* (△). Intensity of the hybridization signals was quantified using an imaging analyzer and were normalized relative to *Actin-1* (Nairn et al., 1988) expression. The *CHS* expression was not shown because the expression level was too low to detect it.

Isolation and characterization of genomic DNA clone

An Arabidopsis genomic library was screened using the full-length CYP73A5 (C4H) cDNA as a probe, yielding a clone consisting of a 291-bp coding region and a 907-bp promoter region of the *CYP73A5* (C4H) gene. The DNA sequence of the coding region was in complete agreement with that of the CYP73A5 (C4H) cDNA (Fig. 9). A putative TATA box was located at 118-bp upstream the ATG translation initiation site (Fig. 9, boxed).

-907	TGTTGTCCTTAAGTTAATCTTAAGATATCTTGAGGTAAATAGAAATAGTTGACTCGTTT	-849
-848	TTATCTTCTTCTTTTTTACCATGAGCAAAAAGATGAAATAAGTTCAAAACGTGACGAA	-789
-788	TCTATATGTTACTACTTAGTATGTGTCAATCATTAAATCGGGAAACTTCATCATTTCAG	-729
-728	GAGTATTACAAAACCTCCTAAGAGTGAGAACGACTACATAGTACATATTTTGATAAAAGAC	-669
-688	TTGAAAACCTTGCTAAAACGAATTTGCGAAAATATAATCATACAAGTGCCAGTGATTTTGA	-609
-608	TCGAATTATTCATAGCTTTGTAGGATGAACCTTAATTAAATAATATCTCACACAAGTATTG	-549
-548	ACAGTAACCTAGTACTATACTATCTATGTTAGAATATGATTATGATATAATTTATCCCCT	-489
-488	CACCTATTCATATGATTTTTGAAGCAACTACTTTCGTTTTTTTAACATTTCTTTTGTG	-429
-428	GTTATTGTTAATGAGCATATTTAGTCGTTTCTTAATTCCTGAAATAGAAAATACAAAG	-369
-368	AGAACTTAGTTAATAGATATGAACATAATCTCACATCCTCCTCCTACCTTCACCAACAC	-309
-308	TTTACATACACTTTGTGGTCTTTCTTTACCTACCACCATCAACAACAACCAAGCCCCA	-249
-248	CTCACACACACGCAATCACGTTAAATTTAACGCCGTTTATTATCTCATCATTACCAACT	-189
-188	CCCACGTACTAACGCCGTTTACCTTGTGCCGTTTGGTCCTCATTCTCTCAAACCAACAAAC	-129
-128	CTCTCCCTCTTATAAAATCCTCTCTCCCTTCTTTATTTCTTCTCAGCAGCTTCTTCTGC	-69
-68	TTTCAATTACTCTCGCCGACGATTTTCTCACCGGAAAAAACAATATCATTGCGGATACA	-9
-8	CAAATAATGGACCTCCTTGTGCTGGAGAAGTCTTTAATCGCCGTCTTCGTGGCGGTG	51
52	ATTCTCGCCACGGTGATTTCAAAGCTCCGCGGCAAGAAATTGAAGCTACCTCCAGGTCCT	111
112	ATACCAATTCCGATCTTCGGAACCTGGCTTCAAGTCGGAGATGATCTCAACCACCGTAAT	171
172	CTCGTCGATTACGCTAAGAAATTCCGGCGATCTCTTCTCCTCCGTATGGGTACAGCGAAC	231
232	CTAGTCGTCGTCTCCTCACCGGATCTAACAAGGAAGTGCTCCTCACTCAAGGCGTTGAG	291

Figure 9. DNA sequence of the TATA-proximal region of the *CYP73A5* gene. The translation initiation codon ATG is located at +1. A putative TATA-box is boxed. Putative cis-acting elements homologous to sequence motifs for boxes P, A, and L (Logemann et al., 1995) are underlined. The deduced amino acid sequence of the coding region is shown below the nucleotide sequence.

In parsley *PAL* gene, three sequence motifs (box P; YTYMMCMAMCMC, box A; CCGTCC, and box L; YCYWACCWACC) have been identified by *in vivo* footprinting as cis-acting regulatory elements (Lois et al., 1989). Logemann et al. (1995) reported that these sequence motifs are conserved among the genes involved in the core reactions of the phenylpropanoid pathway of several plant species including Arabidopsis. To gain an insight into the apparently coordinated-expression of the genes of the phenylpropanoid pathway (Figs. 6, 7, and 8), we compared the 5'-flanking region of the *CYP73A5* (*C4H*) gene with those of the *PAL* genes (Wanner et al., 1995) in Arabidopsis.

Table I. Putative cis-acting elements and their positions relative to the translation initiation ATG codon on the promoters of the *C4H*, *PAL1*, *PAL2*, and *PAL3* genes

	Box P*	Box A*	Box L*
<i>C4H</i>	CcTTCACCAACACt (-321) CTTTACcTACCACC (-285) aTTCACCAACtCCC (-199) CTCaAACCAACCAa (-144)	CCGTtt (-216) CCGTtg (-174) CCGTtt (-160)	CtTTACCTACC (-285) TCaaACCAACC (-143)
<i>PAL1</i>	TcTCAACCAACtCC (-277)	CCGTgt (-395)	gCTTACCTACC (-190)
<i>PAL2</i>	TcTCACCCACCCt (-251)	-	CCTTACCTAaC (-173)
<i>PAL3</i>	-	CCGTCa (-18)	-
Consensus [#]	YTYMMCMAMCMC	CCGTCC	YCYWACCWACC

*The nucleotide sequences which did not match to the consensus sequences were shown by small letters.

[#]Consensus sequences for Box P, A, and L were reported by Logemann et al. (1995) and Y, M, and W indicate C/T, A/C, and A/T, respectively.

Homology analysis was performed in the TATA-proximal promoter region of the *CYP73A5* (*C4H*) gene of Arabidopsis (-1 to -400, relative to the translation initiation codon, ATG), and it was found that there are four sequences homologous to box P, three box A like sequences, and two sequences similar to box L (Fig. 9, underlined and Table I, Logemann et al., 1995). The *PAL1* gene, which was expressed in a manner closely similar to that of *CYP73A5* (Figs. 6, 7, and 8), also contains these three sequence motifs (Logemann et al., 1995). However, on the other hand, box A consensus sequences could not be found in the *PAL2* promoter region (Table I). There are no sequence motifs homologous to box P and L at

least in the promoter region of the *PAL3* gene, of which expression patterns were different from those of *PAL1* and *CYP73A5* (*C4H*). However, it has been reported that the Arabidopsis *PAL3* gene contained two box 1 (box L) sequences (one in the promoter and the other in the second intron) and a box 2 (box P) sequence (Wanner et al., 1995). There is also no AC-rich sequence (box 3, AACCAACAA) in the *CYP73A5* promoter region, which has been suggested as an elicitor-inducible hypersensitive site in bean *CHS15* gene (Ohl et al., 1990).

Discussion

I have demonstrated that *C4H* is encoded by a single copy gene, *CYP73A5*, in Arabidopsis and have examined the *CYP73A5* expression pattern in comparison to other genes coding for enzymes of the phenylpropanoid pathway. It has been reported that biosynthesis of phenylpropanoid-derived natural products is controlled through the modulation of PAL activities, which is achieved in part at the level of transcription (Chappel and Hahlbrock, 1984; Lawton and Lamb, 1987). It is also well known that the *4CL* gene activation is coordinated with certain members of the *PAL* gene family, while only a limited information has been available with respect to *C4H*.

In Arabidopsis, expression patterns of *CYP73A5* (*C4H*), *PAL1*, and *4CL* were closely similar in different tissues, and induction of these genes followed the same time course after the onset of the light period or the wounding treatment, implicating that a common mechanism for their transcriptional activation. Logemann et al. (1995) have recently reported that *C4H* expression in parsley was in concert with the expression of *PAL1/2*, *PAL3*, and both *4CLs*. They found three cis-acting elements (boxes P, A, and L) in most of the *PAL* and *4CL* genes so far reported and suggested that these elements might be involved in coordinating *C4H* gene expression with regulation of *PAL* and *4CL* genes. A similar observation is true in Arabidopsis as well. Specifically, these three cis-acting elements were not only found in the *PAL* and *4CL* genes but also in Arabidopsis *C4H* gene (Table I). This is consistent with the finding in Arabidopsis that *C4H* expression was coordinated with *PAL* and *4CL* in response to wounding and light (Figs. 7 and 8). It has been reported that the activation of the genes of flavonoid biosynthesis involve the Myb and Myc classes of transcription factors, which bind to the core consensus sequence found in boxes P and L (Douglas, 1996).

CHS expression was not dramatically induced by wounding, while the genes involved in the core reactions, *PAL1*, *C4H*, and *4CL*, were. Although *CHS* activity is linked to the core

reactions of general phenylpropanoid pathway (Fig. 1), it is possible that the induction of *PAL*, *C4H*, and *4CL* by wounding does not need to be accompanied by *CHS* expression, which is the first committed step of flavonoid biosynthesis. Similarly, it is possible that the multiple expression patterns of each of the *PAL* isoforms may be explained by different demands in different tissues and conditions for precursors of metabolic pathways branched at *trans*-cinnamic acid, e.g., salicylate (Yalpani et al., 1993; Mauch-Mani and Slusarenko, 1996) and coumarin biosynthesis (Hamerski et al., 1990). In other words, not all *PAL* isozymes may be physiologically linked to *C4H*, although all these isoforms are apparently indistinguishable based on their enzyme-kinetic properties (Appert et al., 1994).

In this chapter, I showed that *PAL*, *C4H*, and *4CL* are coordinately expressed in different tissues and in response to wounding and light. In addition, it was found that the three cis-acting elements in the *PAL* and *4CL* genes (Logemann et al., 1995) are also found in the *C4H* gene. However, Logemann et al. (1995) reported that the three regulatory elements were insufficient in parsley for responding to light or elicitor and that additional elements might be also involved in the regulation of *PAL* gene expression. The analysis of the parsley *4CL-1* gene promoter also suggested that different combinations of cis-acting elements may control expression of *4CL-1* gene in different organs (Hauffe et al., 1993). Further investigation is needed to elucidate the mechanisms operating to control expression of the genes encoding metabolically linked enzymes in such a concerted manner.

Summary

I have isolated a cDNA for a cytochrome P450, cinnamate 4-hydroxylase (*C4H*), of *Arabidopsis thaliana* using a *C4H* cDNA from mung bean as a hybridization probe. The deduced amino acid sequence is 84.7% identical to that of mung bean *C4H*, and the P450 was designated CYP73A5. The CYP73A5 protein was expressed in insect cells using the baculovirus expression system, and the reconstituted *C4H* activity with the recombinant protein (68 nmol/min/nmolP450) confirmed that the CYP73A5 actually encodes the *C4H* of *Arabidopsis*. Southern blot analysis revealed that CYP73A5 is a single copy gene in *Arabidopsis*. A *C4H* promoter region of 907-bp contained all of the three cis-acting elements (boxes P, A, and L) conserved among the *PAL* and *4CL* genes so far reported. In *Arabidopsis*, *C4H* (CYP73A5) expression was apparently coordinated with both *PAL1* and *4CL* in response to the light and wounding, suggesting that these cis-elements are involved in

the regulation of these genes. On the other hand, no significant parallelism was observed among C4H, PAL2, PAL3, and CHS expression patterns, implicating factors supplementary to boxes P, A, and L for transcriptional regulation of these genes.

Chapter IV

Two Isoforms of NADPH-Cytochrome P450 Reductase

in *Arabidopsis thaliana*:

Gene Structure, Heterologous Expression

in Insect Cells, and Differential Expression in *Planta*

Introduction

NADPH-cytochrome P450 reductase (P450-reductase) transfers two electrons from NADPH to P450s. It has been suggested that a single form of P450-reductase is responsible for the electron transport to these diverse P450 isoforms in animals (Porter et al., 1990). In higher plants, P450-reductase proteins and the corresponding cDNAs have been isolated from *Vigna radiata* (Shet et al., 1993) and also from *Catharanthus roseus* (Meijer et al., 1993), in which only a single copy gene for P450-reductase has so far been detected. On the other hand, Benveniste et al. (1991) reported the purification of three isoforms of P450-reductase from the microsomal fractions of *Helianthus tuberosus*. However, it is not clear whether these three reductases were encoded by distinct genes or they were post-translationally modified proteins encoded by a single copy gene. Recently two distinct P450-reductase cDNAs have been isolated from *Arabidopsis thaliana* by Pompon (unpublished results; deposited at GenBank with accession numbers X66016 and X66017, respectively) and also from *Helianthus tuberosus* by Benveniste (unpublished results; deposited at GenBank with accession numbers Z26250 and Z26251, respectively). The results indicate that, in contrast to mammals and yeast, at least some plant species contain a few isoforms of P450-reductase. The occurrence of multiple isoforms of P450-reductase in plants raises questions about physiological roles of the individual P450-reductase isoforms. However, characterization of these P450-reductase isoforms has not been reported yet.

In this chapter, I report the isolation and characterization of the cDNAs and the corresponding genes encoding two isoforms of P450-reductase of *Arabidopsis thaliana*. I demonstrated by expressing the cDNAs in insect cells using a baculovirus expression system that these reductase genes encode functional P450-reductases. Cinnamate 4-hydroxylase activity could be reconstituted at almost the same level in the system containing a

recombinant CYP73A5 (Chapter III, Mizutani et al., 1997) and the recombinant P450-reductases. Genomic organization of the two P450-reductase genes (*AR1* and *AR2*) closely resembled each other with regards to the intron positions and exon sizes, while little similarity was observed in the DNA sequences of their promoter regions. RNA blot analysis showed that *AR1* was constitutively expressed throughout development, whereas *AR2* expression was strongly induced by wounding and light treatments, and the induction time course of *AR2* preceded those of *PAL1* and *CYP73A5*, which encode phenylalanine ammonia-lyase and cinnamate 4-hydroxylase, respectively. Possible regulation mechanisms in coordinating the expression of *AR2* with those of *PAL1* and *CYP73A5* is discussed, on a basis of sequence analyses of the *AR1* and *AR2* promoter regions.

Materials and Methods

Plant materials and treatments

Arabidopsis thaliana ecotype Columbia (Col-0) seedlings were grown as described in Chapter III. Wounding treatment was also performed as described in Chapter III.

For light treatment, 2-week-old plants grown in continuous light were placed in the dark for two days, and then transferred to the light condition for 1 to 12h.

Isolation of cDNA clones of P450-reductases

A cDNA of mung bean P450-reductase was isolated using degenerate oligonucleotide probes designed from the partial amino acid sequences (MR1;RLVAVGLGDDDQ, MR2;LQYGVFGLGNRQYEHFNK, and MR3;LQMDGRYLRDV) determined from a purified mung bean P450-reductase (Chapter I, Mizutani et al., 1993a). A 1.5-kb fragment was obtained by PCR using the set of a sense primer SU18 (5'-CA(A/G)TA(T/C)GA(A/G)CA(T/C)TT(T/C)AA(T/C)AA-3') degenerated from the peptide sequence found QYEHFNK in MR2 and an antisense primer SP69 (5'-TAIC(G/T)ICC(A/G)TCCAT(T/C)-3' from the peptide sequence QMDGRY in MR3. This fragment was used as a hybridization probe to screen a total of 1,000,000 plaques from a mung bean cDNA library under the hybridization conditions as described in Chapter II (Mizutani et al. 1993b). Forty positive clones were isolated and the longest insert was completely sequenced. The cDNA consisted of a 161-bp 5'-untranslated region, a 371-bp 3'-

noncoding region, and a 2073-bp open reading frame encoding a polypeptide of 691 amino acid residues. The deduced protein structure of the cDNA sequence was 30-40% identical to those of mammalian P450-reductases, and the three peptide sequences (MR1, MR2 and MR3) determined from the purified mung bean P450-reductase were found on the predicted amino acid sequence in perfect agreement.

Arabidopsis P450-reductase cDNAs were isolated from a cDNA library prepared from 7-d-old Arabidopsis seedlings (Chapter III, Mizutani et al., 1997) by using the full-length cDNA for mung bean P450-reductase as a probe under low stringency conditions: hybridization, 16h at 50°C in a hybridization buffer containing 1% BSA, 7% SDS, 50 mM sodium phosphate (pH 7.5), and 1 mM EDTA (Church and Gilbert, 1984); washing, 10 min in 6× SSC supplemented with 0.1% SDS at room temperature and 20 min in 2× SSC with 0.1% SDS at 50°C. Twelve positive clones were isolated and divided into two groups according to their partial DNA sequences. The longest clones (AR1 and AR2) from the two groups were completely sequenced.

Isolation of genomic clones for two P450-reductases

Genomic DNA clones were isolated using the full-length cDNAs of AR1 and AR2 as hybridization probes. A total of 100,000 plaques from a λEMBL3 (T7/SP6) library of Arabidopsis ecotype Columbia genomic DNA (Clontech, CA) were screened with either the AR1- or the AR2-probe under high stringency conditions: hybridization, 16h at 65°C in a hybridization buffer described above; washing, 10 min in 2× SSC with 0.1% SDS at room temperature and 30 min in 0.1× SSC containing 0.1% SDS at 65°C. Six positive plaques for each AR1 and AR2 (AR1 clones and AR2 clones) were isolated through two additional rounds of screening, and λDNA was prepared from liquid lysates according to the method of Grossberger (1987). Genomic clones containing the longest inserts were identified by restriction mapping analysis. The inserts were obtained from the AR1 and AR2 clones by digesting with *Xho*I and *Eco*RI, respectively and were subcloned into pCRII vectors (Invitrogen, San Diego, CA) for DNA sequencing.

Heterologous expression in insect cells

The two isoforms of P450-reductase were expressed using a baculovirus expression vector system basically according to the method described as described in Chapter III

(Mizutani et al., 1997).

The expressed P450-reductases were purified from the infected Sf21 cells. Briefly, the infected cells were sonicated and centrifuged at $100,000 \times g$ for 1h. The pellet was homogenized with buffer A containing 20 mM potassium phosphate (pH 7.25), 20% glycerol and 1 mM dithiothreitol, and proteins were solubilized in buffer A supplemented with 1% Emulgen 913 (Kao Atlas, Tokyo, Japan) (buffer B). After centrifugation at $100,000 \times g$ for 1h, the supernatant was applied to a DEAE-Sepharose column (5×10 cm) equilibrated with buffer B, and the protein was eluted with a linear KCl gradient (0-0.5 M) in buffer B. The pooled fractions containing the active P450-reductase were applied to a 2',5'-ADP Sepharose column (1×7 cm) equilibrated with buffer B, and the protein was eluted from the column with 10 mM potassium phosphate buffer (pH 7.7) containing 20% glycerol, 1 mM EDTA, 0.1 mM DTT, and 0.5 mM NADP. The active fractions were further applied to a Mono Q column equilibrated with buffer B, and the column was washed with buffer A until the absorption at 280 nm derived from Emulgen 913 disappeared from the eluted buffer. After the removal of the detergent the protein was eluted with a linear KCl gradient (0-0.5 M) in buffer A.

Assay methods

P450-reductase was assayed by measuring its NADPH-cytochrome *c* reductase activity as described by Imai (1976). The rate of cytochrome *c* reduction was calculated from the absorbance change at 550 nm using an extinction coefficient ($\epsilon = 21 \text{ mM}^{-1} \cdot \text{cm}^{-1}$). C4H activity was reconstituted using the purified recombinant CYP73A5 protein (Chapter III, Mizutani et al., 1997) according to the method described in Chapter I (Mizutani et al., 1993a). The reconstitution system consisted of the recombinant CYP73A5 (5 nM) and the purified recombinant P450-reductases (0.1 unit/mL) in 50 mM potassium phosphate buffer (pH 7.2) containing 0.01% (w/v) sodium cholate, 10 $\mu\text{g/mL}$ dilauroylphosphatidylcholine (Funakoshi, Tokyo, Japan) and 0.4 mM *trans*-cinnamic acid. The reaction was started by adding 0.1 mM NADPH. *p*-Coumaric acid formation was determined by HPLC as described (Chapter I, Mizutani et al., 1993a). Cytochrome P450 was estimated from the CO-difference spectrum (Omura and Sato, 1964).

Preparation and analysis of DNA and RNA

Genomic DNA and total RNA were isolated as described in Chapter III (Mizutani et

al., 1997). For Southern blot analysis, the membrane was hybridized with either the full-length AR1 or AR2 cDNAs as a probe and washed under the high stringency conditions as described in Chapter III. A 282-bp AR1 cDNA fragment was also prepared as a hybridization probe by digesting the AR1 cDNA with *Bam*HI and *Hind*III. Since this probe contained the flavin mononucleotide (FMN) binding domain of P450-reductases (Porter and Kasper, 1986), it was expected to hybridize to any P450-reductase genes of *Arabidopsis* under the low stringency conditions described in Chapter III.

RNA blot analysis, DNA sequencing, and analysis were carried out as described in Chapter III (Mizutani et al., 1997).

Results

Isolation of two different cDNAs encoding P450-reductase of Arabidopsis thaliana

cDNA clones encoding two P450-reductase isoforms (AR1 and AR2) were isolated from an *Arabidopsis* cDNA library using the mung bean P450-reductase cDNA (see Materials and Methods) as a hybridization probe (Fig. 1). The DNA sequences of AR1 and AR2 cDNAs were nearly identical to those of ATR1 and ATR2 cDNAs previously deposited by Pompon at GenBank (accession no. X66016 and X66017, respectively). The small differences (data not shown) might be attributable to the ecotype difference between Columbia (AR1 and AR2) and Landsberg (ATR1 and ATR2).

AR1 cDNA consists of a 2076-bp open reading frame, a 121-bp 5'-untranslated region, a 193-bp noncoding region, and a poly-A tail. The open reading frame encodes a polypeptide of 692 amino acid residues with a calculated molecular mass of 76,765 Da. AR2 cDNA consists of a 2133-bp open reading frame, a 121-bp 5'-untranslated region, a 55-bp noncoding region. Although a stop codon is not found within the 5'-untranslated region upstream the first ATG codon in AR2 cDNA sequence and several methionine residues are seen downstream the first methionine, we assumed the first ATG triplet as the start codon, according to the "first-AUG-rule" in which a first AUG serves as the initiator codon to be used in the translation of about 95% of the eukariotic mRNAs (Kozak, 1987). Thus, the predicted open reading frame of AR2 cDNA encodes a polypeptide of 712 amino acid residues with a calculated molecular mass of 79,124 Da.

A. AR1 cDNA

1	CTCTTCTCTTCGACTCTCTGAATCCTCAAAACCTGGAGAGAGAGAGAGAGAGTACGGGGGAAGAAGATGACTTCTGCTTTGTATGCT	90
1	M T S A L Y A	7
91	TCCGATTGTGTTAAGCAGCTCAAGTCAATTATGGGGACAGATTCGTTATCCGACGATGTTGTACTTGTGATTGCAACGACGCTTTGGCA	180
7	S D L F K Q L K S I M G T D S L S D D V V L V I A T T S L A	37
181	CTAGTAGCTGGATTGTGGTGTGTTATGGAAGAAAACGCGCGGATCGGAGCGGGAGCTGAAGCCTTTGATGATCCCTAAGTCTCTT	270
37	L V A G F V V L L W K K T T A D R S G E L K P L M I P K S L	67
271	ATGGCTAAGGACGAGGATGATGATTGGATTGGGATCCGGGAAGACTAGAGTCTCTATCTTCTGGGTACGCAGACTGGAACAGCTGAG	360
67	M A K D E D D D L D L G S G K T R V S I F F G T Q T G T A E	97
361	<u>GGATTGTCTAAGTCATTATCCGAAGAAATCAAAAGCGAGATATGAAAAGCAGCAGTCAAGTCATTGACTTGGATGACTATGCTGCGGAT</u>	450
97	G F A K A L S E E I K A R Y E K A A V K V I D L D D Y A A D	127
451	<u>GATGACCACTATGAAGAGAAATGAAGAAGGAACTTTGGCATTTTCTGTGTTGCTACTTATGAGATGGAGAGCCTACTGACAAATGCT</u>	540
127	D D Q Y E E K L K K E T L A F F C V A T Y G D G E P T D N A	157
541	<u>GCCAGATTTTACAAATGGTTTACGGAGCAAAATGAACGGGATATAAGCTTCAACAACCTAGCATATGGTGTGTTGCTCTTGGTAATCGC</u>	630
157	A R F Y K W F T E E N E R D I K L O Q L A Y G V F A L G N R	187
631	CAATATGAACATTTAATAAGATCGGGATAGTCTTGTGATGAAGATTATGTAAGAAGGTGCAAGCGTCTTATTGAAGTCGGTCTAGGA	720
187	O Y E H F N K I G I V L D E E L C K K G A K R L I E V G L G	217
721	GATGATGATCAGAGCATTGAGGATGATTTAATGCCTGAAAGAATCACTATGGTCTGAGCTAGACAAGCTCCTCAAAGACGAGGATGAT	810
217	D D D Q S I E D D F N A W K E S L W S E L D K L L K D E D D	247
811	AAAAGTGTGGCACTCCTTATACAGCTGTTATTCTGAATACCGGGTGGTGAATCATGCTCGCTTACAACCTCAAAATCAATGGAA	900
247	K S V A T P Y T A V I P E Y R V V T H D P R F T T Q K S M E	277
901	TCAATGTGGCAATGGAATACTACTATTGACATTCATCCTCGCAGTTGATGTTGCTGTGAGAGGAGCTTACACACATGAA	990
277	S N V A N G N T T I D I H H P C R V D V A V Q K E L H T H E	307
991	TCTGATCGGTCTTGCAATTCATCTCGACTTCGACATATCCAGGACGGGTATTATATGAAACAGGTGACCATGTAGGTGTATATGCTGAA	1080
307	S D R S C I H L E F D I S R T G I T Y E T G D H V G V Y A E	337
1081	AATCATGTTGAGATAGTTGAAGAAGCTGGAAAATGGCTGGCCACTCTTTAGATTTAGTATTTCCATACATGCTGACAGGAAGATGGC	1170
337	N H V E I V E E A C K L L G H S L D L V F S I H A D K E D G	367
1171	TCCCATTTGGAAGCGCAGTCCCGCTCCTTCCCTGGTCCATGCACACTGGGACTGGTTTGGCAAGATACGCAGACCTTTTGAACCTT	1260
367	S P L E S A V P P P F P G P C T L G T G L A R Y A D L L N P	397
1261	CCTCGAAATCTCGCTTAGTTGCCTTGGCGGCTATGCCACTGAACCAAGTGAAGCCGAGAACTTAAGCACCTGACATCACTGATGGA	1350
397	P R K S A L V A L A A Y A T E P S E A E K L K H L T S P D G	427
1351	AAGCATGACTACTCACAATGGATTGTTGCAAGTCAGAGAAGCTTTTAGAGGTGATGGCTGCTTTTCCATCTGCAAAACCCCACTAGGT	1440
427	K D E Y S Q W I V A S Q R S L L E V M A A F P S A K P P L G	457
1441	GTATTTTGTGCTGAATAGCTCTCTGCTACAACTCGTTACTACTCCATCTCATCTGCAAGATTTGGGGCCCAAGTAGAGTTTCATGTT	1530
457	V F F A A I A P R L Q P R Y Y S I S S C Q D W A P S R V H V	487
1531	ACATCCGCACTAGTATATGGTCCAACCTCTACTGGTGAATCCACAAGGGTGTGTGTTCTACGTGGATGAACATGCAGTCTCTGCGGAG	1620
487	T S A L V Y G P T P T G R I H K G V C S T W M K N A V P A E	517
1621	AAAAGTCATGAATGATGGAGCCCAATCTTTATTCGAGCATCTAATTTCAAGTTACCATCCAACCTTCAACTCCAATCGTTATGCTG	1710
517	K S H E C S G A P I F I R A S N F K L P S N P S T P I V M V	547
1711	GGACCTGGGACTGGGCTGGCACCTTTTAGAGGTTTCTGCAGGAAGGATGGCACTAAAGAAGATGGAGAAGAAGTGGTTCATCTTTG	1800
547	G P G T G L A P F R G F L Q E R M A L K E D G E E L G S S L	577
1801	CTCTCTTTGGGTGTAGAAATCGACAGATCTACTTTATATACGAGGATGAGCTCAATAATTTTGTGATCAAGGCGTAAATCTGAGCTC	1890
577	L F F G C R N R Q M D F I Y E D E L N N F V D Q G V I S E L	607
1891	ATCATGGCATCTCCCGTGAAGGAGCTCAGAAGGAGTATGTTCAACATAAGATGATGGAGAAAGCAGCACAAGTTTGGGATCTAATAAAG	1980
607	I M A F S R E G A Q K E Y V Q H K M M E K A A Q V W D L I K	637
1981	GAAGAAGGATATCTATGATGCGGTGATGCTAAGGGCATGGCGAGGGACGTCACCCGAACTCTACACACCATTTCTCAGGAGCAAGAA	2070
637	E E G Y L Y V C G D A K G M A R D V H R T L H T I V Q E Q E	667
2071	GGTGTGAGTCTGTCAGAGGAGGCTATAGTTAAGAACTTCAAAACGAAGAGATACCTCAGAGATGTCGTTGATGATGATGA	2160
667	G V S S S E A E A I V K K L Q T E G R Y L R D V W *	692
2161	TTTATTGTAATCTCTCTGCTTCTCAACTTTTCTTGGTACGCTGCTTACCATTTAGATATTTATTGTTTATTCTCTCTGCTCTC	2250
2251	CCTTCTTTTAGTCAAAACATGAATGATTGTCAAAGTTTCTTTGTATGAACAAAAGTATAAACTACAAATAAATCTGTGAGAGTGATAA	2340
2341	AA	2342

Figure 1. Nucleotide and deduced amino acid sequences of the two Arabidopsis P450-reductase cDNAs.

A: AR1 cDNA; B: AR2 cDNA. The AR1 cDNA fragment which was used as a probe for genomic Southern blot analysis (Fig. 4B) is underlined. The putative polyadenylation signal is double underlined. Intron positions are indicated by vertical lines with the number of the position from the N-terminus.

	2	TCACCTACACCACACCTAACCAATCCCCCTACGATTACACAGAGAGAGATCTTCTCTCTCCTTCTCTCTCTCTCTCTT	91	
	1			1
	92	TCTTCTTCTAGCTACAACATCTACAACGCCATGTCCCTCTTCTTCTTCTCGTCAACCTCCATGCATCATGGCAGCAATCATCAA	181	
	1	M S S S S S S S S T S M I D L M A A I I K	20	
	182	GGAGAGCCTGTAAATTGTCTCGGACCCAGCTAATGCCCTCCGCTTACGAGTCCGCTAGCTGCTGAATTATCCTCTATGCTTATAGAGAATCGT	271	
	20	G E P V I V S D P A N A S A Y E S V A A E L S S M L I E N R	50	
	272	CAATTCCGCATGATTGTTACCACTTCCATTGCTGTTCTTATTGGTTCATCGTTATGCTCGTTGGAGGAGATCCCGCTTCTGGGAATCA	361	
	50	O F A M I V T T S I A V L I G C I V M L V W R R S G S G N S	80	
	362	AAACCTGTCGAGCCTCTTAAGCCTTTGGTTATTAAGCCTCGTGAGGAGAGATGATGATGGGCGTAAGAAAAGTTACCATCTTTTTCGGT	451	
	80	K R V E P L K P L V I K P R E E E I D D G R K K V T I F F G	110	
	452	ACACAAACTGGTACTGCTGAAGGTTTTCGAAGGCTTTAGGAGAAGAAGCTAAAGCAAGATATGAAAGACCAGATTCAAAATCGTTGAT	541	
	110	T Q T G T A E G F A K A L G E E A K A R Y E K T R F K I V D	140	
	542	TTGATGATTACGCGGCTGATGATGATGAGTATGAGGAGAAATTGAAGAAAGAGGATGGCTTTTCTTCTTACCCACATATGGAGAT	631	
	140	L D D Y A A D D D E Y E E K L K K E D V A F F F L A T Y G D	170	
	632	GGTGAGCCTACCGACAATGCGCGAGATTCTACAAATGGTTACCGAGTGGGAATGACAGAGGAGATGGCTTAAGAAGTGAAGTATGGA	721	
	170	G E P T D N A A R F Y K W F T E G N D R G E W L K N L K Y G	200	
	722	GTGTTTGGATTAGGAACACAGCAATATGAGCATTTTAATAAGGTTGCCAAAGTTGTAGATGACATTTCTGTGSAACAGTGCACAGCGT	811	
	200	V F G L G N R O Y E H F N K V A K V V D D I L V E Q G A Q R	230	
	812	CTTGATCAAGTTGGTCTTGGAGATGATGACCAAGTATTGAAGATGACTTTACCGCTTCGCGAGAAGCATTGTGGCCCGAGCTTGATACA	901	
	230	L V Q V G L G D D D O C I E D D F T A W R E A L W P E L D T	260	
	902	ATACTGAGGGAAGAGGGGATACAGCTGTTGCCACACCATACACTGCAGCTGTTTAGAATACAGAGTTTCTATTACGAGCTCTGAAGAT	991	
	260	I L R E E G D T A V A T P Y T A A V L E Y R V S I H D S E D	290	
	992	GCCAAATKCAATGATATAAACATCGGCAATGGGAATGGTTACACTGTGTITGAGCTCAACATCCCTACAAAGCAAAATGCGCTGTAAAA	1081	
	290	A K F N D I N M A N G N G Y T V F D A Q H P Y K A N V A V K	320	
	1082	AGGGAGCTTCATACTCCCAGTCTGATCGTTCTTGATCCATTGGAAATTTGACATTGCTGGAAGTGAGCTTATCTATGAAACTGCGAGAT	1171	
	320	R E L H T P E S D R S C I H L E F D I A G S G L T Y E T G D	350	
	1172	CATGTTGGTGTACTTTGTGATAACTTAACTGAAACTGTAGATGAAGCTCTTAGATTGCTGGATATGTCACCTGATACTTATTCTCACCT	1261	
	350	H V G V L C D N L S E T V D E A L R L L L D M S P D T Y F S L	380	
	1262	CACGCTGAAAAGAAGACGGCACCCAATCAGCAGCTCACTGCCTCCTCCCTTCCACCTTGAACCTGAGAACAGCGCTTACACGATAT	1351	
	380	H A E K E D G T P I S S S L P P P F P P C N L R T A L T R Y	410	
	1352	GCATGTCTTTTGTAGTTCTCCAAAGAAGCTGCTGTTAGTTGCGTTGGCTGCTCATGCATCTGATCCTACCGAAGCAGAAGCATTAAACAC	1441	
	410	A C L L S S P K K S A L V A L A A H A S D P T E A E R L K H	440	
	1442	CTTGCTTCACTGCTGGAAGGATGAATATTCAAAGTGGGTAGTAGAGTCAAAGAGTCTACTTGGAGTGATGGCCGAGTTTCTTCA	1531	
	440	L A S P A G K D E Y S K W V V E S Q R S L L E V M A E F P S	470	
	1532	GCCAAGCCACCATTGGTGTCTTCTTCGCTGGAGTTGCTCCAAGGTTGACGCTAGGTTCTATTGATATCATCATGCCCAAGATTGCT	1621	
	470	A X P P L G V F F A G V A P R L Q P R F Y S I S S S P K I A	500	
	1622	GAAACTAGAATTACGTCACATGTGCATCGTGTATGAGAAAATGCCAACTGGCAGGATTGATAAGGAGTGCTGTCCACTTGGATGAAG	1711	
	500	E T R I H V T C A L V Y E K M P T G R I H K G V C S T W M K	530	
	1712	AATGCTGTGCCTTACGAGAAGAGTGAAACCTGTTCTCGGCGCCGATATTGTTAGGCAATCCAACCTCAAGCTTCTCTGATTCTAAG	1801	
	530	N A V P Y E K S E N C S S A P I F V R Q S N F K L P S D S K	560	
	1802	GTACCGATCATCATGATCGGTCCAGGGACTGGATTAGCTCCATTACAGAGGATTCTTCAGGAAAAGACTAGCGTTGGTAGAATCTGGTGT	1891	
	560	V P I I M I G P G T G L A P F R G F L Q E R L A L V E S G V	590	
	1892	GAACCTGGGCCATCAGTTTGTGTTCTTTGGATGAGAAACCGTAGAATGATTTCATCTACGAGGAAGAGCTCCAGCGATTGTTGAGAGT	1981	
	590	E L G P S V L F F G C R N R R M D F I Y E E E E L Q R F V E S	620	
	1982	GGTGCTCTCGCAGAGCTAAGTGTCCCTTCTCTCGTGAAGGACCCACCAAAGATACGTACAGCACAAGATGATGGACAAGCTTCTTGAT	2071	
	620	G A L A E L S V A F S R E G P T K E Y V Q H K M M D K A S D	650	
	2072	ATCTGGAATATGATCTCTCAAGGAGCTTATTATATGTITGTGGTGACGCCAAAGGATGGCAAGAGATGTTACAGATCTCTCCACACA	2161	
	650	I W N M I S Q G A Y L Y V C G D A K G M A R D V H R S L H T	680	
	2162	ATAGCTCAAGAACAAGGCTCAATGGATTCAACTAAAGCAGAGGGCTTCGTGAAGAATCTGCAAAAGAGTGAAGATATCTTAGAGATGTA	2251	
	680	I A Q E Q G S M D S T K A E G F V K N L Q T S G R Y L R D V	710	
	2252	TGGTAACGAAACTATTGAAGCCACACTCACTGTGACTTATATTATATATATACGGCACAGAAATGCCACATTATGATGATCATTA	2341	
	710	W *	711	
	2342	AGTTTGTGATCGCAAGAAGAAAGAACTCCTTTTPTTTCATTTTAAATTCTTTTCATATTATTITTGACAACCTATTTTTTTAACTC	2431	
	2432	TTGTTATATATCCCCCACCCTAAGTAAAGAAAAATGCATAAAAAAAAA	2480	

Fig. 2 shows the alignment of the deduced primary structures of AR1 and AR2 proteins with those of the reductases from other plant species. P450-reductases from animals and yeast consist of several functional domains such as hydrophobic membrane anchoring region, FMN-, FAD-, NADPH-binding regions and also the binding sites for P450 and cytochrome *c* (Porter and Kasper, 1986; Nishimoto, 1986, Yabusaki et al, 1988). The amino acid sequences of these functional domains were well conserved between the two Arabidopsis reductases (Fig. 2), irrespective of less structural similarity in the N-terminal region. The deduced primary structures of AR1 and AR2 proteins are 63% identical to each other (Table I). Interestingly, AR1 protein is more similar to the reductases of *P. aureus* and *V. sativa* than to AR2 (Table I). On the other hand, AR2 protein is 73% identical to the *C. roseus* reductase and is 75% and 70% identical to HTR1 and HTR2 from *H.tuberosus*, respectively (Table I).

Table I. Identity of amino acid sequences of Arabidopsis P450- reductases to those of the reductases from various species

	AR1	AR2	Accession No.
	Identity (%)		
<i>Arabidopsis thaliana</i> (AR1)	-	63	in this study
<i>Arabidopsis thaliana</i> (AR2)	63	-	in this study
<i>Helianthus tuberosus</i> (HTR1)	67	75	Z26250
<i>Helianthus tuberosus</i> (HTR2)*	65	70	Z26251
<i>Vicia sativa</i>	74	67	Z26252
<i>Phaseolus aureus</i>	74	67	in this study
<i>Catharanthus roseus</i>	63	73	X69791
<i>Homo sapiens</i>	41	40	S90469
<i>Rattus norvegicus</i>	40	41	M12516
<i>Drosophila melanogaster</i>	41	42	X93090
<i>Caenorhabditis elegans</i>	37	39	U21322
<i>Saccharomyces cerevisiae</i>	37	37	D13788
<i>Aspergillus niger</i>	38	38	Z26938
<i>Bacillus megaterium</i> *	34	33	J04832

* HTR2 is a partial cDNA clone lacking the N-terminal region (Fig. 2)

* In *B. megaterium*, a gene encodes P450-BM-3 which consists of two domains corresponding to cytochrome P450 and P450-reductase (Ruettinger et al., 1989), and therefore, the region coding for the P450-reductase domain was used for the homology analysis.

Particularly, the N-terminal structures of AR1, *P. aureus*, and *V. sativa* reductases were shorter than those of AR2, *C.roseus*, and *H.tuberosus* 1 reductases (Fig.2).

AR1	MSSSSSSS	HIDLMAAIK	GEPIVSDPA	NASAYESVAA	ELSSMLIENR	OFANIVTTSI	AVLIGCIVHL	VNRSGSGNS	K--R-VEPLK	PLVIKPREE	96
AR2	MAS---	NSD	LVRAYESFLG	V-SIGDSVSD	-LLIATTTA	AVVIG-LLVF	LWKKSS-DRS	KEVKPVVVR	DLM-HEEEEE	72	
<i>P. aureus</i>	MTSS-	NSD	LVRTIESALG	I-SIGDSVSD	-VVIATTTA	AVIIG-LLVF	LWKKSP-DRS	RELKPVIVPK	FTV-KHEDD	73	
<i>V. sativa</i>	MDSSEKLS	P-ELMSATLK	GAKLDGNS-	-SDSGVAVSP	AVHAKLENK	ELVHILTTSV	AVLIGCVVVL	IWRSSSGS-G	K--KVVEPK	LIVKSVVEP	94
<i>C. roseus</i>	MQPETVETP	FAVL---NNK	---VIDNGNS-	-----THP	VLTMLLDNQ	-LLVLTSA	AVLLGFIVVF	VNRSSSGKKS	G--KELEPK	IVVPRQVE	82
<i>H. tuberosus1</i>											
<i>H. tuberosus2</i>											
FMN-pyrophosphate											
AR1	LDLGSGRTRV	SIFPGTGTGT	ADGFAKALSE	EIKARYEKAA	VKVIDLDYYA	ADDDQYEELK	KKETLAPFCV	ATTGDGEPTD	NAARFYKWT	EE-NERDIKL	174
AR2	-EIDDGRKKV	TIFFPGTGTGT	ADGFAKALGE	EAKARYEKTR	FKIVOLDYYA	ADDDQYEELK	KKEDVAFFFL	ATTGDGEPTD	NAARFYKWT	EG-NDRGEML	194
<i>P. aureus</i>	VDVAAGRIKV	TIFFPGTGTGT	ADGFAKALAE	EIKARYEKAA	VKVVDLDYYA	ADDDQYEELK	KKESLVFFHL	ATTGDGEPTD	NAARFYKWT	EGKDERGIML	172
<i>V. sativa</i>	VEVDGRGRIKV	TVFYGTGTGT	ADGFAKALAE	EIKARYEKAV	VKVVDLDYYA	IDDDQYEELK	KKETLVFFHL	ATTGDGEPTD	NAARFYKWT	EGKEERGIML	173
<i>C. roseus</i>	EEIDEGKKKF	TIFFPGTGTGT	ADGFAKALAE	EAKARYEKAV	IKVIDLDYYA	ADDDQYEELK	KKETLAPFLL	ATTGDGEPTD	NAARFYKWT	EG-NDRGIML	193
<i>H. tuberosus1</i>	QEVDDGKKKV	TIFFPGTGTGT	ADGFAKALVE	EAK-RYEKAV	FKVIDLDYYA	ADDDQYEELK	KKESLAPFLL	ATTGDGEPTD	NAARFYKWT	EG-DOKGIML	180
<i>H. tuberosus2</i>	---DGKKKV	TVFYGTGTGT	ADGFAKALVE	EAKARYEKAV	FKVIDLDYYA	ADDDQYEELK	KKESLAPFLL	ATTGDGEPTD	NAARFYKWT	EG-EPKIML	97
P450, cytochrome c											
AR1	QQLAYGVFAL	GNRQYEFNK	IGIVLDEELC	KKGAKRLIEV	GLGDDDDQIE	DDFNARKESL	WSELQKLLKD	EDD-KSVATP	YTAVIPEYRV	VTHDPRFT--	271
AR2	KNLYKGVFGL	GNRQYEFNK	VAKVVDLILV	EQAQRILVGV	GLGDDDDQIE	DDFTAWREAL	WPELDLILRE	EGD-TAVATP	YTAALVLEYRV	SIHSDADAKF	293
<i>P. aureus</i>	QQLTYGVFGL	GNRQYEFNK	IGKIVDEELA	EQAQRILVGV	GLGDDDDQIE	DDFSANKESL	WPELDQLLRD	EDDANTVSTP	YTAALVLEYRV	VTHDPRFT--	270
<i>V. sativa</i>	QQLTYGVFGL	GNRQYEFNK	IGKIVDEELT	EQAQRILVGV	GLGDDDDQIE	DDFNARKESL	WPELDQLLRD	EDDVNTASTP	YTAALVLEYRV	VTHDPRFT--	271
<i>C. roseus</i>	KNLYKGVFGL	GNRQYEFNK	IAKVVDKVA	EQAQRILVPL	VLGDDDDQIE	DDFAAWRENV	WPELDNLLRD	EDD-TTVSTT	YTAALVLEYRV	VFPDKSDGLI	292
<i>H. tuberosus1</i>	NKLHYGVFGL	GNRQYEFNK	IAKVVDKGLT	EQAQRILVFPV	GLGDDDDQIE	DDFSANKESL	WPELDQLLRD	EDD-KTAATP	YTAALVLEYRV	VFPDKSDGLI	278
<i>H. tuberosus2</i>	NKLHYGVFGL	GNRQYEFNK	IAKVVDKGLV	EQAQRILVFPV	GLGDDDDQIE	DDFTAWRENV	WPELDQLLRD	ODD-TSVATP	YTAALVLEYRV	VFPDKSDGLI	195
FAD-pyrophosphate											
AR1	--TQKSHESN	VANGNTTIDI	HHPCRVDVAV	QKELHTHESD	RSCIHLEFDI	SRTGITTYETG	DHVGVAENH	VEIVEEAGKL	LGHSLDLVPS	IHADKEDGSP	369
AR2	NDIMHANGNG	Y----TVFDA	GHFYKANVAV	KRELHTPESD	RSCIHLEFDI	AGSGITYETG	DHVGVLCDNL	SETVDEALRL	LDMSPDIYFS	LBAKEDGTP	389
<i>P. aureus</i>	--STYDNESN	VANGNTTIDI	HHPCRVDVAV	QKELHTPESD	RSCIHLEFDI	SGTISITYETG	DHVGVAENH	NETVEETGKL	LQGNLDLFFS	LHTDKEDGTS	368
<i>V. sativa</i>	--PSYENRHN	VANGNTTIDI	HHPCRVDVAV	QKELHTPESD	RSCIHLEFDI	SGTISITYETG	DHVGVAENH	DETVEEAGKL	LQGNLDLFFS	LHTDKEDGTS	369
<i>C. roseus</i>	SEAN-CHANG	YANGNTTIDI	HHPCRVDVAV	QKELHTPESD	RSCIHLEFDI	AGTGLSYETG	DHVGVAENH	SETVEEAGKL	LNLPPPTTFS	LHADKEDGTP	391
<i>H. tuberosus1</i>	SENH-SOTNG	---HTVHDA	HHPCRVDVAV	QKELHTPESD	RSCIHLEFDI	SHTGLSYETG	DHVGVAENH	IEVVEEAGKL	IGLPADTYFS	LHIDNEDGTP	371
<i>H. tuberosus2</i>	-DEN-SOTNG	---HAYHDA	HHPCRVDVAV	QKELHTPESD	RSCIHLEFDI	SHTGLSYETG	DHVGVAENH	SEVVEEAGKL	IGLPADTYFS	LHIDNEDGTP	289
FAD-isalloxazine-ring											
AR1	LESA-VPPPF	PGPCTLTGTL	ARYADLLNPF	RKSALVALAA	YATEPSEAEK	LKHLTSPQCK	DETSQWIVAS	QRLLEVMMA	FPSAKPPLGV	FFAAIAPRLQ	468
AR2	IS-SSLPPPF	P-PCNLRTAL	TRYACLLSSP	RKSALVALAA	HASDPTAEER	LKHLTSPQCK	DETSQWIVAS	QRLLEVMMA	FPSAKPPLGV	FFAAIAPRLQ	487
<i>P. aureus</i>	LGGS-LLPPF	PGPCTLTGTL	ARYADLLNPF	RKAALLALAT	HASEPSEAEK	LKFLSPQCK	DETSQWIVAS	QRLLEVMMA	FPSAKPPLGV	FFAAIAPRLQ	467
<i>V. sativa</i>	LGGS-LLPPF	PGPCTLTGTL	ARYADLLNPF	RKAALLALAT	HASEPSEAEK	LKFLSPQCK	DETSQWIVAS	QRLLEVMMA	FPSAKPPLGV	FFAAIAPRLQ	468
<i>C. roseus</i>	LGGSLLPPPF	P-PCNLRTAL	TRYADLLNPF	RKSALVALAA	YASDPNEADR	LKFLSPQCK	DETAQSLVAN	QRLLEVMMA	FPSAKPPLGV	FFAAIAPRLQ	490
<i>H. tuberosus1</i>	LGGSLLPPPF	P-PCNLRTAL	TRYADLLNPF	RKSALVALAA	HASDPTAEER	LKFLSPQCK	DETAQSLVAN	QRLLEVMMA	FPSAKPPLGV	FFAAIAPRLQ	472
<i>H. tuberosus2</i>	LGGSLLPPPF	P-PCNLRTAL	TRYADLLNPF	RKSALVALAA	HASDPTAEER	LKFLSPQCK	DETSQWIVAS	QRLLEVMMA	FPSAKPPLGV	FFAAIAPRLQ	388
NADPH-ribose, pyrophosphate											
AR1	PRYYSIBSSQ	DWAPSRVHVT	SALVYGPPTP	GRIHKGVCST	WMKNAVPAEK	SHDCSGAPIF	IRASHFKLPS	NPSTPIVHVG	PGTGLAPFRG	FLQERNALKE	568
AR2	PRYYSIBSSP	KIAETRIHVT	CALVYEKPTP	GRIHKGVCST	WMKNAVPAEK	SHDCSGAPIF	IRASHFKLPS	DSKVPIMIG	PGTGLAPFRG	FLQERNALKE	587
<i>P. aureus</i>	PRYYSIBSSP	RFAPQVRHVT	CALVYGPPTP	GRIHKGVCST	WMKNAVPAEK	SHDCSGAPIF	IRASHFKLPS	DHSIPIMIG	PGTGLAPFRG	FLQERNALKE	567
<i>V. sativa</i>	PRYYSIBSSP	RFAPQVRHVT	CALVYGPPTP	GRIHKGVCST	WMKNAVPAEK	SHDCSGAPIF	IRASHFKLPS	DHSIPIMIG	PGTGLAPFRG	FLQERNALKE	568
<i>C. roseus</i>	PRYYSIBSSP	RMAPSRVHVT	CALVYEKPTP	GRIHKGVCST	WMKNAVPAEK	SHDCSGAPIF	IRASHFKLPS	DPKVPVIMIG	PGTGLAPFRG	FLQERNALKE	590
<i>H. tuberosus1</i>	PRYYSIBSSP	RMAPSRVHVT	CALVYEKPTP	GRIHKGVCST	WMKNAVPAEK	SHDCSGAPIF	IRASHFKLPS	DPKVPVIMIG	PGTGLAPFRG	FLQERNALKE	572
<i>H. tuberosus2</i>	PRYYSIBSSP	RMAPSRVHVT	CALVYEKPTP	GRIHKGVCST	WMKNAVPAEK	SHDCSGAPIF	IRASHFKLPS	DPKVPVIMIG	PGTGLAPFRG	FLQERNALKE	488
NADPH-nicotinamide											
AR1	DGEELGSSLL	FFGCRNRQMD	FIYEDELNPF	VDQGVISELI	HAPSREGAOK	ETVQKNNK	AAQVNDLKE	EGYLYVCGDA	KGMARDVHRT	LBTIVQEQG	668
AR2	SGVELGSPVL	FFGCRNRQMD	FIYEDELNPF	VDQGVISELI	HAPSREGAOK	ETVQKNNK	AAQVNDLKE	GATLYVCGDA	KGMARDVHRT	LBTIVQEQG	687
<i>P. aureus</i>	DGVOIGPALL	FFGCRNRQMD	FIYEDELNPF	VDQGVISELI	HAPSREGAOK	ETVQKNNK	AAQVNDLKE	GGYLYVCGDA	KGMARDVHRT	LBTIVQEQG	667
<i>V. sativa</i>	DGVOIGPALL	FFGCRNRQMD	FIYEDELNPF	VDQGVISELI	HAPSREGAOK	ETVQKNNK	AAQVNDLKE	GGYLYVCGDA	KGMARDVHRT	LBTIVQEQG	668
<i>C. roseus</i>	EGAEELGAVF	FFGCRNRQMD	FIYEDELNPF	VDQGVISELI	HAPSREGAOK	ETVQKNNK	AAQVNDLKE	GATLYVCGDA	KGMARDVHRT	LBTIVQEQG	690
<i>H. tuberosus1</i>	SGVELGSPVL	FFGCRNRQMD	FIYEDELNPF	VDQGVISELI	HAPSREGAOK	ETVQKNNK	AAQVNDLKE	GATLYVCGDA	KGMARDVHRT	LBTIVQEQG	672
<i>H. tuberosus2</i>	AGTELGSLL	FFGCRNRQMD	FIYEDELNPF	VDQGVISELI	HAPSREGAOK	ETVQKNNK	AAQVNDLKE	GATLYVCGDA	KGMARDVHRT	LBTIVQEQG	588
AR1	VSSSEAEIV	KKLQSGRYL	RDVN	692							
AR2	NDSTKAEIV	KKLQSGRYL	RDVN	711							
<i>P. aureus</i>	VDSTKAEIV	KKLQSGRYL	RDVN	691							
<i>V. sativa</i>	ADSKAEATV	KKLQSGRYL	RDVN	692							
<i>C. roseus</i>	NDSTKAEIV	KKLQSGRYL	RDVN	714							
<i>H. tuberosus1</i>	LDSTKAEIV	KKLQSGRYL	RDVN	696							
<i>H. tuberosus2</i>	LDSTKAEIV	KKLQSGRYL	RDVN	612							

Figure 2. Amino acid alignment of plant P450-reductases. The deduced amino acid sequences of AR1 and AR2 are aligned with those of P450-reductases from *P. aureus*, *V. sativa*, *C. roseus*, *H. tuberosus* by the program, Clustal W (Thompson et al., 1994), and the alignment was further modified to maximize the sequence homology. Identical amino acid residues are shadowed, and dashes are inserted to maximize the sequence homology. The putative functional regions (involved in the interaction with FMN, FAD, NADPH, P450, and cytochrome c) are indicated above the sequences by arrow.

Thus, plant P450-reductases can be apparently divided into two groups: AR1-type; AR1 and the P450-reductases of *P. aureus* and *V. sativa*, and AR2-type; AR2, the *C. roseus* P450-reductase and the two *H. tuberosus* reductase (HTR1 and HTR2). In contrast to the high homology among plant P450-reductases, AR1 and AR2 proteins were only 32-41% identical to the P450-reductases from other organisms (Table I), suggesting that the sequence variation among the plant P450-reductases has occurred after the divergence from other organisms.

Biochemical properties of AR1 and AR2 proteins expressed in insect cells

I expressed the AR1 and AR2 cDNAs in insect cells using a baculovirus expression vector system. SDS-PAGE analysis (Fig. 3) showed that new intense bands of 78 kDa and 80 kDa appeared in the microsomal fraction of the insect cells upon infection with the recombinant AR1-virus and the AR2-virus, respectively.

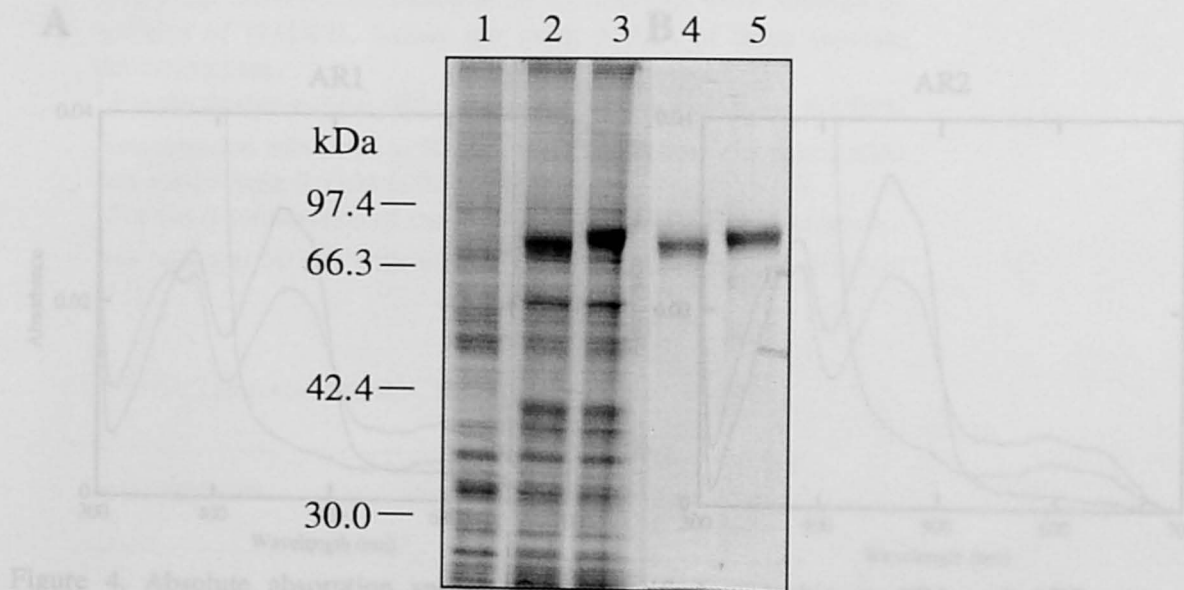


Figure 3. Synthesis of recombinant AR1 and AR2 proteins. SDS-PAGE was performed using 10% polyacrylamide slab gel, and proteins were visualized with Coomassie brilliant blue R-250. Lane 1, Solubilized fraction of microsomes of mock-infected Sf21 cells (10 mg of protein); lane 2, solubilized fraction of microsomes of Sf21 cells (10 mg of protein) infected with the recombinant virus containing AR1 cDNA; lane 3, solubilized fraction of microsomes of Sf21 cells (10 mg of protein) infected with the recombinant virus containing AR2 cDNA; lane 4, purified AR1 protein (200 μ g of protein); lane 5, purified AR2 protein (200 μ g of protein). The migration of size standard is shown to the left of the gel.

These proteins were not found in the mock-infected cells (Fig. 3, lane 1). The difference in the apparent molecular masses of these newly appeared proteins reflected the molecular masses of AR1 (76,765 Da) and AR2 (79,124 Da) calculated from their deduced primary

structures. The microsomal NADPH-cytochrome *c* reductase activity of the recombinant virus infected cells was 1000-fold higher than that of the mock-infected cells. These recombinant AR1 and AR2 proteins were purified to homogeneity by a three-step column chromatography (Fig. 3, lanes 4 and 5, respectively).

The recombinant AR1 and AR2 proteins showed closely similar absolute absorption spectra characteristic of flavoproteins (Fig. 4A and B). Oxidized forms showed prominent peaks at 455 and 380 nm, typical of a flavoprotein, and the 455 nm peak disappeared when the enzymes were fully reduced by sodium dithionite. Aerobic treatment of the reductases with 50 μ M NADPH produced a spectrum having a broad peak around 600 nm, which is typical for the stable semiquinone form of flavoproteins. These spectral properties of the recombinant AR1 and AR2 proteins are very similar to that of the P450-reductase purified from rabbit microsomes (Iyanagi and Mason, 1973).

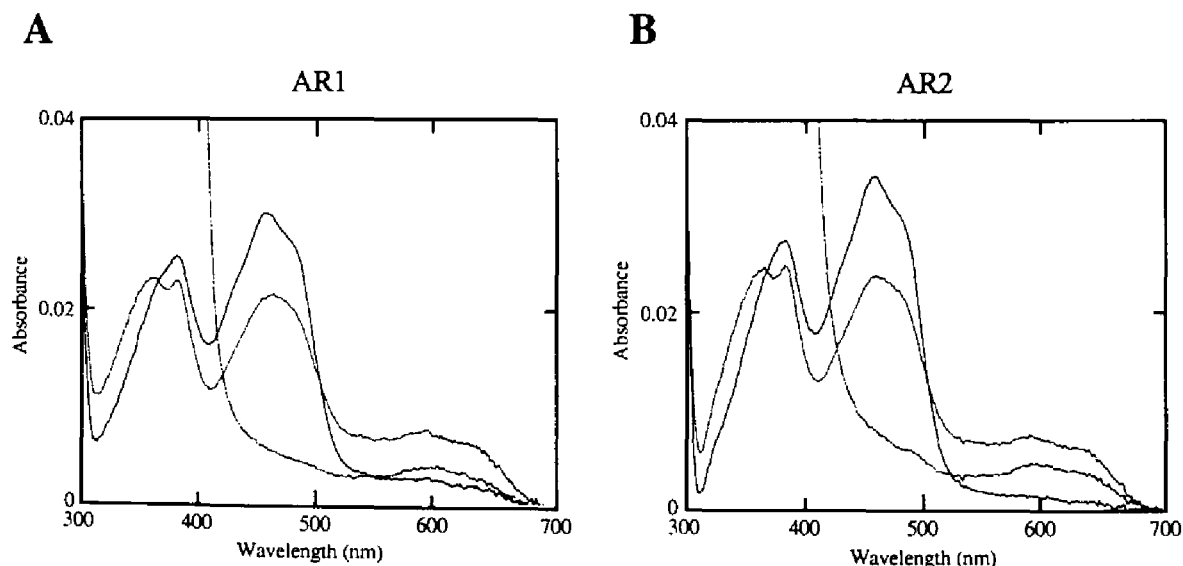


Figure 4. Absolute absorption spectra of the purified recombinant AR1 and AR2 proteins. A, the purified recombinant AR1 protein; B, the purified recombinant AR2. The one-electron reduced semiquinone forms were prepared by adding 1 mM NADPH to a final concentration of 25 μ M, and the spectra were recorded after incubating for 10 min at 25 °C. A few grains of sodium dithionite were added to completely reduce the reductases. —, Oxidized form; , semiquinone form; - - - - - , completely reduced form.

The recombinant AR1 and AR2 proteins were also indistinguishable in terms of the K_m values for NADPH and cytochrome *c* (Table II), which were comparable to those of the reductases from *P. aureus* (22.1 μ M and 24.8 μ M for NADPH and cytochrome *c*, respectively) and the *H. tuberosus* reductases (Benveniste et al., 1989). In the reconstitution system with either the AR1 or AR2 protein, the recombinant CYP73A5 protein (Mizutani et al, 1997) was able to

catalyze the cinnamate 4-hydroxylase reaction at the same rate (Table. III), indicating that AR1 is as efficient as AR2 in donating electrons to CYP73A5. However, it is still possible that these two P450 reductases have different specificity towards individual P450s.

Table II. The K_m values for cytochrome *c* and NADPH

P450-reductase	K_m^a	
	cytochrome <i>c</i> ^b	NADPH ^c
	μM	
AR1	24.3 \pm 2.3	21.9 \pm 1.4
AR2	22.5 \pm 1.8	23.0 \pm 2.9

^a Reactions were carried out in 0.3 M potassium phosphate (pH 7.7). After 2min-preincubation at 28°C, reactions were initiated by addition of NADPH. Values are mean \pm S.D. of three separate determinations.

^b For the determination of the K_m value for cytochrome *c*, NADPH concentration was fixed at 50 μM , and cytochrome *c* concentration was varied from 0.5 μM to 300 μM .

^c For the determination of the K_m value for NADPH, cytochrome *c* was added at 50 μM with varied NADPH concentrations (0.5-300 μM).

Table III. The reconstituted CYP73A5 activity

Assay condition	cinnamate 4-hydroxylase activity ^b
	$nmol/min/nmol$ P450
Complete ^a AR1	63.6 \pm 4.3
AR2	70.2 \pm 5.6
- NADPH	n.d. ^c
- P450-reductase	n.d. ^c
- CYP73A5	n.d. ^c

^a Complete reaction mixture contained 50 mM potassium phosphate (pH 7.25), 5 nM recombinant CYP73A5, 0.1 unit/ml recombinant P450-reductase (AR1 or AR2), 0.01% (w/v) sodium cholate, 10 $\mu g/ml$ dilauroylphosphatidylcholine, 0.1 mM NADPH, and 0.4 mM *trans*-cinnamic acid.

^b Reactions were carried out in the complete reaction mixture at 30°C and were initiated by addition of NADPH. Values are mean \pm S.D. of three separate determinations.

^c not detected.

Genomic Southern blot analysis

Genomic DNA was digested with *EcoRI*, *HindIII*, or *XbaI*, and hybridization was performed at the high stringency conditions (65 °C) using either the full-length cDNAs for AR1 or AR2 as a probe (Figs. 5A and C). Both *EcoRI* and *XbaI* digestion do not cut the AR1 cDNA, and a single hybridization signal was observed when hybridized with the AR1 probe (Fig. 3A). Three fragments were found in the digestion with *HindIII* (Fig. 5A), which has two restriction sites in the AR1 cDNA sequence. There are two *EcoRI* sites, a *HindIII* site, and no *XbaI* site in the AR2 cDNA sequence. Therefore three bands with the *EcoRI* digestion, two close bands with *HindIII*, and a single band with *XbaI* digestion were detected using the AR2 probe (Fig. 5C). The results indicated that each of the AR1 and AR2 cDNA probes hybridized with an independent single gene and did not cross-hybridize each other under the high stringency conditions.

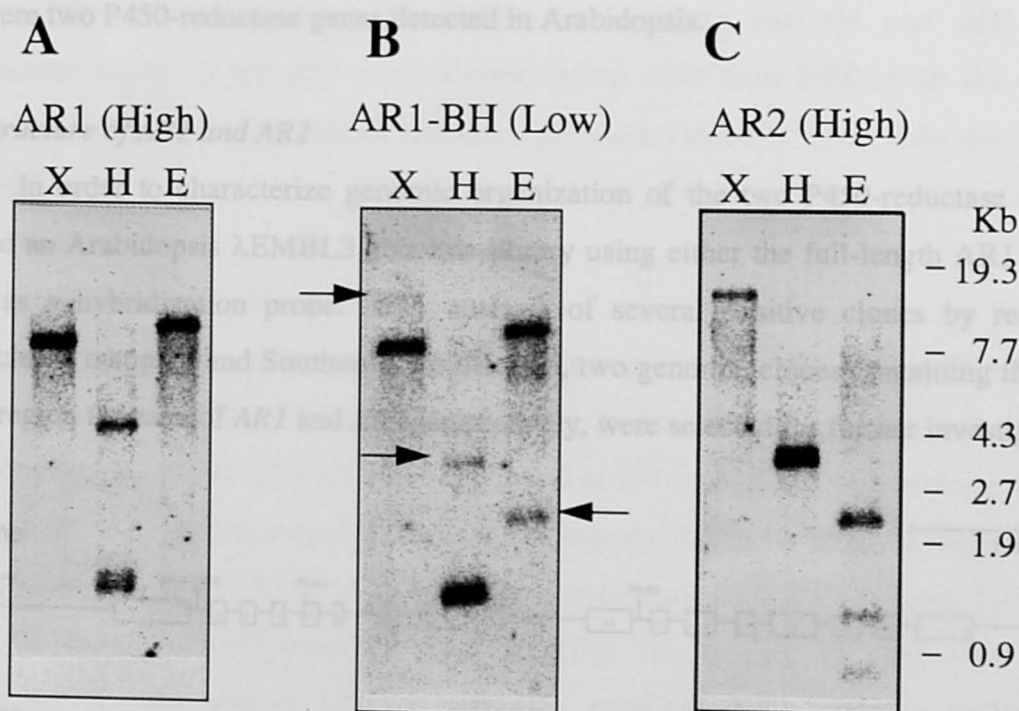


Figure 5. Southern blot analysis of the P450-reductase genes. *A.thaliana* Columbia genomic DNA (1 µg) was digested with the restriction enzymes; X, *XbaI*; H, *HindIII*; E, *EcoRI*. The digested DNA was separated on 0.7% agarose gel, blotted onto a nylon membrane, and hybridized with a ³²P-labeled probe. A, the full-length AR1 cDNA was used as a probe at the high stringency; B, the AR1 cDNA fragment obtained by digestion with *Bam*HI and *Hind*III (Fig.1A, underlined) was used as a probe at the low stringency. The weak hybridization bands, which could be ascribed to the AR2 gene were indicated by arrows; C, the full-length AR2 cDNA was used as a probe under the high stringency. The migration of size marker is shown to the right of the blots.

A 282-bp fragment was obtained from the AR1 cDNA by digesting with *Bam*HI and *Hind*III (Fig. 1A, underlined) and was used as a probe under the low stringency conditions (50 °C) (Fig. 3B). This 282-bp fragment could serve as a universal probe to detect any P450-reductase genes in Arabidopsis (74.5% identity between AR1 and AR2 at the DNA level), since it encodes a domain containing the putative FMN-binding site conserved in all P450-reductases so far reported (Fig. 2; Porter, 1991). In each digestion, two gene fragments were observed at strong and weak intensities, 8-kb and 15-kb fragments in the *Eco*RI digestion, 1.5-kb and 3.8-kb fragments in *Hind*III, and 11-kb and 2.4-kb fragments in *Xba*I digestion, respectively (Fig. 5B). When the hybridization patterns were compared (Figs. 5A, B, and C), these strong and weak hybridization signals detected under the low stringency conditions could be ascribed to *AR1* and *AR2*, respectively. In other words, the 282-bp fragment from the AR1 cDNA hybridized with only those DNA fragments derived from *AR1* and *AR2*. Thus, there were two P450-reductase genes detected in Arabidopsis.

Gene structure of *AR1* and *AR2*

In order to characterize genomic organization of the two P450-reductase genes, I screened an Arabidopsis λ EMBL3 genomic library using either the full-length AR1 or AR2 cDNA as a hybridization probe. After analysis of several positive clones by restriction endonuclease mapping and Southern hybridization, two genomic clones containing the entire coding region for each of *AR1* and *AR2*, respectively, were selected for further investigation.

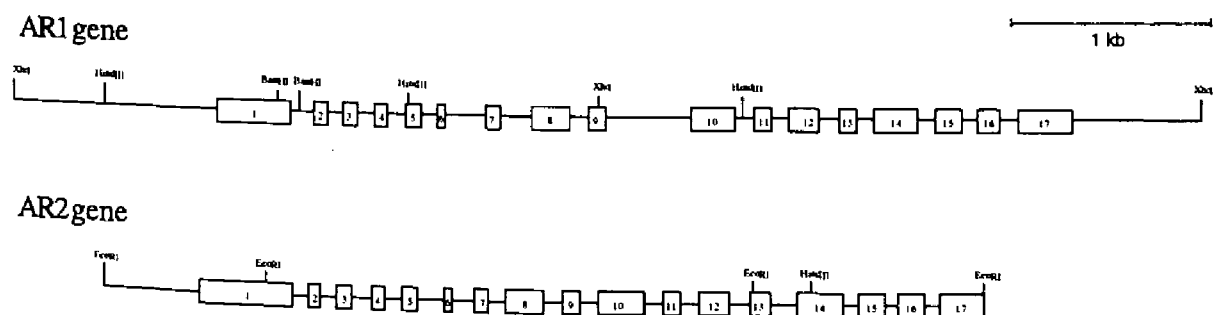


Figure 6. Gene organization of the *AR1* and *AR2* genes. Exons are indicated by boxes 1-17, and introns are indicated by solid lines.

The physical maps of *AR1* and *AR2* gene organizations are shown in Fig. 6. Sequence comparison of *AR1* and *AR2* with those of the corresponding cDNAs showed that

both *AR1* and *AR2* contain 17 exons and 16 introns (Figs. 1 and 6). All of the exon-intron boundaries are consistent with the proposed junction sequence 'gt...ag' rule (Hanley and Schuler, 1988). Interestingly, the *AR1* and *AR2* coding sequences were divided exactly at the same positions by introns (Fig. 1). This extensive conservation of intron placements suggest that the two reductase genes might have evolved by way of the duplication of a common ancestral gene. In contrast to the striking conservation of the intron positions, lengths of the corresponding introns vary between *AR1* and *AR2*, and little sequence similarity was observed (data not shown). Furthermore, there was a correlation between the exon organization and the functional domains in both *AR1* and *AR2* (Figs. 1 and 2). Porter et al. (1990) proposed that the exon organization of the rat reductase gene correlated with the functional domains of the reductase. Comparison of the gene structures of Arabidopsis reductases with that of rat reductase showed that 3 intron positions (intron 9, 11, and 12) were consistent on their aligned amino acid sequences (data no shown).

Additional finding from the P450-reductase gene analysis was that the 3'-nontranslated region of the *AR1* gene is overlapping more than 2000-bp in the opposite direction with the *CER2* gene, which encodes a protein involved in wax production (Negruk et al., 1996).

The promoter regions of AR1 and AR2

The DNA sequences of the 5'-flanking regions of *AR1* and *AR2* were also determined as shown in Fig. 7A and B. In a 1154-bp promoter region of *AR1*, both a putative TATA box and a putative CCAAT box were found at 309-bp and 348-bp upstream the ATG translation initiation codon, respectively (Fig. 7A, boxed). A 616-bp promoter region of *AR2* also contained a putative TATA box at 171-bp and two putative CCAAT boxes at 208-bp and 237-bp upstream the ATG codon, respectively (Fig. 7B, boxed). O'Leary et al. (1994) reported that the promoter of rat P450-reductase gene possesses neither a TATA nor a CCAAT box but contain GC-rich consensus sequences for the transcription factor Sp1 and is similar to those of housekeeping genes. In contrast to the rat reductase promoter, both the TATA and CCAAT boxes but no Sp1 consensus sequences were found in the *AR1* and *AR2* promoters.

- 1153	GATCCGACAGACAGAATAAACACCATTTGTAAGAGCTGGTGGCTGGAAGATTGCAGTGCCT	- 1034
- 1093	CATCTTCGTTTCGATAAAGTGTGGTCTTAATTCCTCAACAAATGTGCCCATTTCTCTGGCAG	- 1034
- 1033	TAACATCTAGTTCTGACATTGCTTCATCTTTTTTTTTAGGACATGGCTGCTCATTCAAGGC	- 974
- 973	TGTGACTTTGAAGTGGCAAGCAGAAAACTGGCGCTGGCAAAATATCGCCAGGTGATC	- 914
- 913	TGGCAAGGAAGGGTCTTAAATTCACAGTTCACAGATGAGGAGCTAAGAAACAAAGTTCGC	- 854
- 853	TGTCGGTTGACACCAGAAGAGATGGGTTTACTGCTGTCAGCTTTGGGTTTCAGTAACAAT	- 794
- 793	ACCCGCTCTCTATCTAGCGTCACACCAAGGTTTACGGAGGAGAAGCAAGAATCTCAACGCTG	- 734
- 733	CGTAAACTGTTCCCGGGATTGAAAACAAGAAAAGCCTAGCCTCTGCTGAGGAACTAGCT	- 674
- 673	GATGTACAAGGAAAGCTTCTCTGATGGCTGCAGTCGATTACTACGTGAGTATGAAGAGC	- 614
- 613	GACATCTTCATCTCTGCATCTCCCGGGAACATGCACAACGCATTACAGGGCGCATAGAGCC	- 554
- 553	TACTTGAATCTCAAAACCATAAGACCCAACATGATACTGTTAGGACAAGTCTTTGTGAAC	- 494
- 493	AAGAGCTTAGATTGGTCAGAGTTTGAAGGAGCAGTAATGAATGGGCACAAAACAGGCCAA	- 434
- 433	GGCCAGCTTCGCTACGCAAAACAGAAACAGTCTATTACACTTATCCAGCTCCTGATTGC	- 374
- 373	ATGTGTAAAGTAGCTTAGCATTCCCTTTTACTGTCTCAATTTTTTGTGATTTTTTAA	- 314
- 313	AATTTTAAATATATGTTTATGTAATCGATGATTAGTAAATTAATTGAGGAATCATAT	- 254
- 253	TATGTACACTAAATTCAGTTTAGTAAAAAATGGATTAAATGTGATTACTTATTGACAA	- 194
- 193	AGGTAAAAATAAAGTCAAATAAGAAAAGGAGAGAAGGACCCACTCAGATTCAACGCCCCA	- 134
- 133	CCACCGATAACTCAATTAACCTCTCTCTCTTGGCTTCTGCAACGCTGAGCCACCATCGTC	- 74
- 73	CTCTTCTCTCTCTCTTCGACTCTCCTGAATCCTCAAACCTGGAGAGAGAGAGAGAGAGT	- 14
- 13	GACGGGGGAAGAAGATGACTTCTGCTTTGTATGCTTCCGATTTTGTAAAGCAGCTCAAGT	46
1	M T S A L Y A S D L F K O L K	16

-615	AAC TATTTTATTTT <u>TACCCCAACAAG</u>	GTTAAATCTTACTTTTTTCTATTTTGCTCTTAATTT	-556
	P		
-555	CCCATTCTGATTAAAGAACCAATCTCATATACTTTTCTATTTTTGTTAGACATGAATTC		-496
-495	AAATATATATCTTACATTTTATAAAATTAATTTTCAATCAAAAATCTGTAGTATTATACA		-436
-435	<u>ACCGTGAT</u> GTTCCTAAAAAATACCTGTCATACAGCTGTTATGCTTCTTAAAAATAAGATAA		-376
	A		
-375	AATTATATTTACTGTATTAAAGCTAGATTGTCAAAAATTAACAGCAGAATAAACGTAA		-316
-315	TTAACCATCGTTAATATCAAACGGTTAGGTGCGGTTTGACCGGTCGTGTTTGGTGAGCGC		-256
-255	TGGTAATAATTTCAACCAT <u>CCAAT</u> CAATCACAGACAAAATTGAAACCC <u>CCAAT</u> TAATAAT		-196
	P.L. A		
-195	ACATAAAATGAAAAACATGATGAAT <u>TATAAA</u> ATAAAACTCTGTTCTGTCTCAAAAAACTC		-136
-135	AAAAGGAGTCTCTCTCACCTACACCACACCTAACCAATCCCCCTACGATTACACAGA		-76
-75	GAGAGATCTTCTTCTCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTAGCTA		-16
-15	CAACATCTACAACGCCATGTCTCTTCTTCTTCTTCTCGTCAACCTCCATGATCGATCTCAT		44
1	M S S S S S S S S T S M I D L		14

Figure 7. The nucleotide sequences of the TATA-proximal regions of the *AR1* and *AR2* genes. A, the *AR1* promoter region; B, the *AR2* promoter region. The translation initiation codon ATG is located at +1. A putative TATA box and a CAAT box are boxed. Putative cis-acting elements homologous to sequence motifs for boxes P, A, and L (Logmann et al., 1995; Chapter III, Mizutani et al., 1997) are underlined. The deduced amino acid sequences of the coding regions are shown below the nucleotide sequences.

As described in Chapter III, the promoter of *CYP73A5* encoding cinnamate 4-hydroxylase contains three sequence motifs (boxes P, A, and L) conserved among the genes for phenylalanine ammonia-lyase (*PAL*) and 4-coumarate CoA: ligase (*4CL*) involved in general phenylpropanoid pathway (Mizutani et al., 1997). These elements are thought to be important in controlling the coordinated-expression of these genes under different environmental conditions (Logemann et al., 1995). It should be noted that, in the *AR2* promoter region, there were two sequences similar to box P and box A, and a sequence homologous to box L (Fig. 7B, underlined). On the other hand, the *AR1* promoter contained only two box A like sequences but no sequences homologous to boxes P and L (Fig. 7A, underlined).

Expression manners of AR1 and AR2 in Arabidopsis

RNA gel blot hybridization was performed to study expression manners of *AR1* and *AR2*, using each of the full-length *AR1* and *AR2* cDNAs as gene specific probes. The organ-specific and age-dependent expression patterns of the two reductases are shown in Fig. 8.

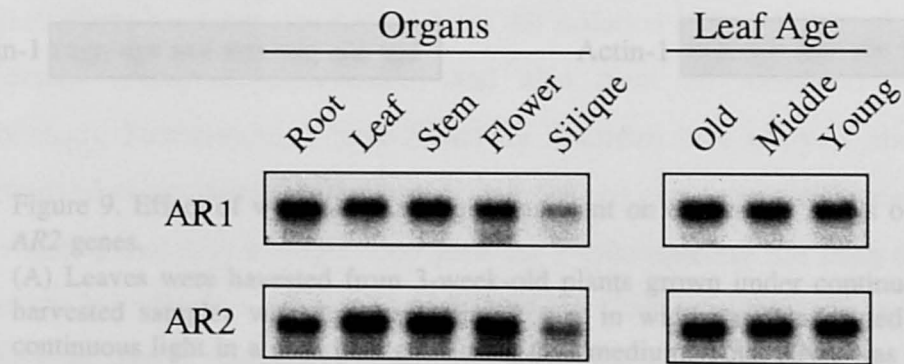


Figure 8. Tissue specific expression of the *AR1* and *AR2* genes in Arabidopsis.
 (A) Total RNA was isolated from the roots and leaves of 3-week-old plants, from inflorescence stems and flowers of 4-week-old plants, and from the siliques of 5-week-old plants.
 (B) Total RNA was isolated from the leaves of different age of 3-week-old plants with 12 leaves. The first and second leaves represented the older leaves. The middle-aged leaves were collected from the fourth and fifth positions, and the younger leaves were from the ninth and tenth positions counted from the bottom. Plants were grown under continuous light. Five µg of total RNA was separated on formaldehyde agarose gels, transferred to nylon membranes, and hybridized to the indicated probes.

The *AR1* mRNA was about 1.3-fold abundant in roots and stems than in leaves and flowers. On the other hand, the *AR2* expression was 2-fold higher in leaves, stems, and flowers than in

roots. Both the P450-reductase genes showed the least expression in siliques. Strong age-dependency was not found in the expression levels of *AR1* and *AR2*. This result contrasts with the results in Chapter V that most of P450s isolated from *Arabidopsis* were more highly expressed in older leaves than in younger leaves (Mizutani et al., submitted).

Expression manners of *AR1* and *AR2* in response to wounding and light treatments were also investigated (Fig. 9A and B), comparing with those of phenylalanine ammonia-lyase (*PAL1*) and cinnamate 4-hydroxylase (*CYP73A5*).

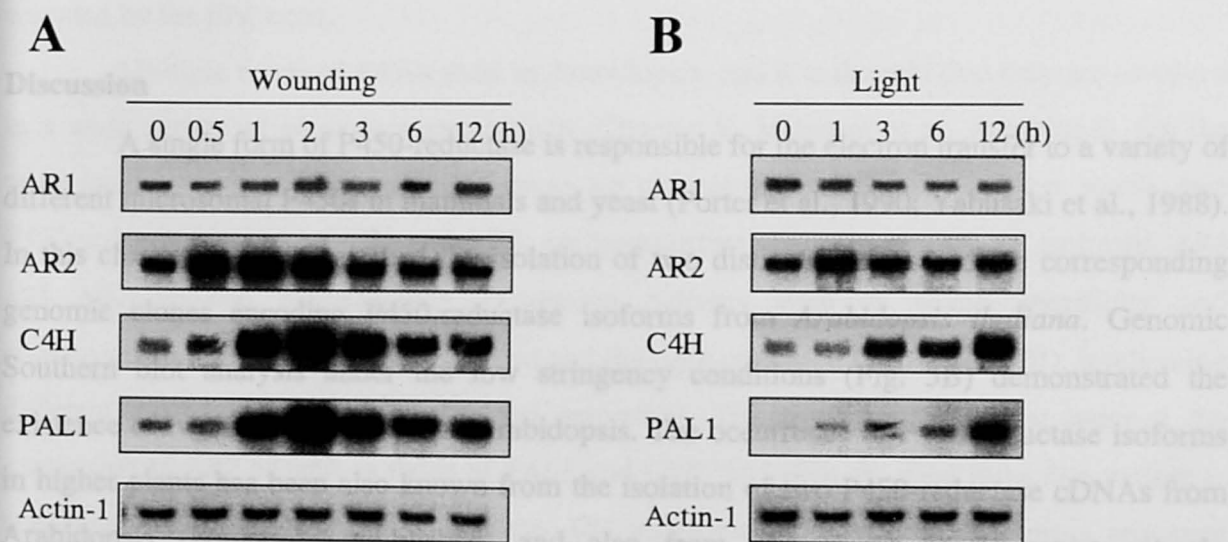


Figure 9. Effect of wounding and light treatment on expression levels of the *AR1* and *AR2* genes.

(A) Leaves were harvested from 3-week-old plants grown under continuous light. The harvested samples were cut into strips 2 mm in width, and incubated for 9h under continuous light in a petri dish containing GM medium. Total RNA was isolated at the times indicated after wounding and analysed by RNA gel blotting (5 µg per lane) using the probes indicated.

(B) Two week-old plants grown under continuous light were placed in the dark for two days and returned to the light condition. Total RNA was isolated from the leaves at times indicated after the onset of light period (0-9h) and analysed by RNA gel blotting (5 µg per lane) using the probes indicated.

The strong and coordinated-induction of *PAL1* and *CYP73A5* by wounding (Fig. 9A) was consistent with my previous observation (Chapter III, Mizutani et al., 1997). The *AR2* expression level was significantly enhanced by wounding, and it reached a maximum (4-fold) within 1h of the treatment and then gradually decreased to a basal level. On the other hand, the *AR1* expression did not change during the treatment. The expression levels of *PAL1* and

CYP73A5 continually increased to a 10-fold within 12h of the light treatment (Fig. 9B). The *AR2* expression also increased 3-fold as fast as 1h of the onset of the light period. However, the induction was transient and the *AR2* mRNA level gradually decreased to a basal level. On the other hand, the *AR1* expression slightly decreased with light treatment. To sum up, *AR1* was constitutively expressed, while *AR2* was induced in response to the wounding treatment and light. The *AR2* induction was followed by those of *PAL1* and *CYP73A5*, of which the magnitude of induction was far greater than that of *AR2*.

Discussion

A single form of P450-reductase is responsible for the electron transfer to a variety of different microsomal P450s in mammals and yeast (Porter et al., 1990; Yabusaki et al., 1988). In this chapter, I have described the isolation of two distinct cDNAs and the corresponding genomic clones encoding P450-reductase isoforms from *Arabidopsis thaliana*. Genomic Southern blot analysis under the low stringency conditions (Fig. 5B) demonstrated the existence of two reductase genes in *Arabidopsis*. The occurrence of P450-reductase isoforms in higher plants has been also known from the isolation of two P450-reductase cDNAs from *Arabidopsis* (Pompon, unpublished) and also from *H. tuberosus* (Benveniste et al., unpublished). Furthermore, I have found by Southern blot analysis under low stringency conditions (data not shown) a few copies of P450-reductase gene in *P. aureus*, whereas it has been reported that only a single copy gene for P450-reductase has been detected under high stringency hybridization conditions in *V. radiata* (Shet et al., 1993) and also in *C. roseus* (Meijer et al., 1993). Benveniste et al. (1991) reported that two or three proteins were detected in the microsomes from all higher plants they tested by Western blot analysis using polyclonal antibodies prepared against the *H. tuberosus* reductase. I cannot rule out the possibility that some of these immunoreactive proteins might be post-translationally modified proteins encoded by a single copy gene. In fact, there are potential glycosylation sites in both *AR1* protein (284-N) and *AR2* protein (31-N and 358-N). Nonetheless, available results suggest that the occurrence of the P450-reductase isoforms may be common in higher plants.

I have also investigated the gene organization of *AR1* and *AR2* (Fig. 6). The highly conserved exon organization implicates that the two genes have evolved from a common ancestral gene. The most striking differences were observed in the region encoded by the first

exon. Specifically, the N-terminal portion encompassing the signal anchor sequence of AR2 is much longer than that of AR1, and the amino acid sequences in this region are also very different between them. On the other hand, the sequence similarity was 68% between AR1 and AR2 proteins when calculated without the domain encoded by the first exon. Homology analysis of higher plant P450-reductases indicated that plant P450-reductases may be divided into AR1- and AR2-types. For example, the P450 reductases from mung bean and periwinkle could be grouped into AR1- and AR2-type, respectively. It should be noted that the structural differences between these two types of P450-reductases were found specifically in the region encoded by the first exon.

Multiple forms of P450s exist in Arabidopsis, and it is thought that they are involved in a wide variety of secondary metabolism (Chapter V, Mizutani et al., submitted). On the other hand, only two P450-reductase genes are present in Arabidopsis. In other words, two P450-reductase isoforms, AR1 and AR2, are transporting electrons to all P450 isoforms. *AR1* and *AR2* showed the different expression patterns such as tissue specificity and responsiveness to the wounding treatment and the light condition (Fig. 8A and B), implicating different mechanisms for transcriptional activation. The wounding and light induced the expression of *AR2* as well as those of *PAL1* and *CYP73A5* (Fig. 9). This induction pattern is consistent with the finding that only the *AR2* promoter contained the consensus sequence motifs for the putative cis-acting elements (boxes P, A, and L) (Fig. 7B) involved in the coordinated-expression of *PAL* and *4CL* genes (Logmann et al., 1995). Inversely, the promoter region of the constitutive P450-reductase, *AR1*, did not contain the sequence motifs for boxes P and L (Fig. 7A). I have presented the evidence that *CYP73A5* in Arabidopsis is also regulated together with the other phenylpropanoid pathway genes (i.e., *PAL1* and *4CL*) via a mechanism containing these three cis-acting elements (Chapter III, Mizutani et al., 1997). It has been also reported that *PAL1* (Ohl et al., 1990) and *CYP73A5* (Bell-Long et al., 1997) were strongly expressed in the vascular tissues of roots and leaves of Arabidopsis. Both *PAL1* and *CYP73A5* contain the putative cis-elements, which were also found in the *AR2* promoter. Whereas it has not been fully elucidated whether these cis-elements are involved in the tissue specific expression, it is possible that *AR2* expression may be tissue-specifically coordinated with those of *PAL1* and *CYP73A5* and participate in the dynamic fluctuation of the activity of the phenylpropanoid pathway. On the other hand, the different time course of induction between the expression of *AR2* and those of *PAL1* and *CYP73A5* suggested another

mechanism in controlling the *AR2* mRNA levels in *Arabidopsis*.

I have shown that not only *CYP73A5* but also several other P450 genes were induced by either wounding or light (Chapter V, Mizutani et al., submitted), indicating that the total number of P450 molecules in plants should fluctuate in response to changes of environmental conditions. It has been proposed that about 10-15 molecules of P450s are interacting with one P450-reductase molecule in the microsomal membrane (Klinger, 1990), and the P450-reductase level should also vary in keeping an appropriate P450/P450-reductase ratio. It is therefore possible that plant P450-reductases should be inductibly expressed in a manner coordinated with inducible P450s. However, it remains unclear why both the inducible- and constitutive type of P450-reductases with indistinguishable enzymatic properties coexist specifically in higher plants. It is possible that the constitutive type (*AR1*) may be necessary to keep a basal P450-reductase level and that the expression of the inducible type (*AR2*) may synchronize with the fluctuation of the total P450 level, which should change dynamically in response to environmental conditions. The K_m values for NADPH and cytochrome *c* were closely similar between the two P450-reductase isoforms (Table II), which could couple with *CYP73A5* protein at essentially the same efficiency in the reconstitution system (Table III). However, it is not necessarily true that *AR1* and *AR2* proteins may have similar affinity to individual P450s. Specifically, the different structures of the N-terminal portions between *AR1* and *AR2* (Fig. 2) may be implying distinct physiological roles for each of the P450-reductase isoforms. Preparation of specific antibodies and promoter analysis of *AR1* and *AR2* could constitute interesting approaches for further investigation of physiological roles of the two reductases. Also, the inter- and intracellular localization of *AR1* and *AR2* should be studied together with those of various P450s *in planta*. The production of transgenic plants, specifically loss-of-function mutants, of *AR1* and *AR2* may also help to understand the relationship between the reductases and certain P450-dependent reactions.

Summary

I have investigated two NADPH-cytochrome P450 reductase (P450-reductase) isoforms encoded by separate genes (*AR1* and *AR2*) in *Arabidopsis thaliana*. First, we isolated *AR1* and *AR2* cDNAs using a mung bean P450-reductase cDNA as a probe. The recombinant *AR1* and *AR2* proteins produced using a baculovirus expression system showed

similar K_m values for cytochrome *c* and NADPH, respectively. In the reconstitution system with a recombinant CYP73A5, the recombinant AR1 and AR2 proteins gave the same level of cinnamate 4-hydroxylase activity (about 70 nmol/min/nmol P450). The AR2 gene expression was transiently induced by 4- and 3-fold within 1h of wounding and light treatments, respectively, and the induction time course preceded those of *PAL1* and *CYP73A5*. On the contrary, the *AR1* expression level did not change during the treatments. Analysis of the *AR1* and *AR2* gene structure revealed that the AR2 promoter specifically contained three putative sequence motifs (boxes P, A, and L), which are involved in the coordinated-expression of *CYP73A5* and other phenylpropanoid pathway genes. These results implicated the possibility that *AR2* might be metabolically linked to the induced levels of phenylpropanoid pathway enzymes.

Chapter V

Cytochrome P450 Superfamily in *Arabidopsis thaliana*:

cDNAs Isolation, Differential Expression, and RFLP Mapping of Multiple Cytochromes P450

Introduction

The observation of various P450-dependent reactions *in vitro* suggested existence of multiple diverse P450s in a single plant species (Donaldson and Luster DG, 1991; Bolwell et al., 1994). While several plant P450 cDNAs have been isolated, the number of unique P450 genes and the variety of physiological roles in a single plant species has not been established yet. Recent progress in molecular biology has facilitated the isolation of novel P450 clones from plants. In fact, several cDNAs belonging to three P450 families could have been isolated from eggplant using an eggplant CYP75 cDNA as a probe (Umemoto et al., 1993; Toguri et al., 1993), four P450 cDNAs of the CYP71A subfamily have been also cloned from maize (Frey et al., 1995), as have. Another successful approach is based on PCR using primers designed from conserved amino acid sequence around the heme-binding domain of P450 proteins. This strategy is particularly useful to isolate cDNAs of low abundance, or instable P450s. Holton et al. (1993), employing this approach, have isolated 18 P450 cDNA fragments from petunia and have found that two clones encoded P450s with flavonoid 3',5'-hydroxylase activity. It has also been reported that 15 cDNA fragments were isolated from periwinkle with a PCR approach (Meijer et al., 1993), and 4 P450 cDNAs were isolated from pea using reverse transcription-PCR (Frank et al., 1996). However, it is usually difficult to elucidate physiological functions of the P450s isolated through these molecular biological approaches.

In this chapter, I report the isolation of 13 full-length and 3 partial P450 cDNAs from *Arabidopsis thaliana*. 11 P450s have been assigned to 5 distinct families, CYP71B, CYP73A, CYP76C, CYP83, and CYP91. The N-terminal structures of the deduced proteins showed properties similar to those of mammalian microsomal P450 proteins, suggesting microsomal localization of these *Arabidopsis* P450s. RNA blot analysis revealed tissue characteristic expression patterns for each of the P450s, which responded differently to

wounding and light. One of the isolated P450s, CYP73A5, has been already identified as cinnamate 4-hydroxylase (Chapter III, Mizutani et al., 1997), whereas physiological functions of the other 15 P450s are still unknown. As a first step towards the elucidation of their physiological functions, genetic loci of four P450 genes have been determined by RFLP mapping, and a linkage map with the loci of known mutations is reported.

Materials and Methods

Plant materials/treatments

Arabidopsis thaliana ecotype Columbia (Col-0) and ecotype Landsberg *erecta* (Lehle Seeds) seedlings were grown as described in Chapter III.

Wounding and dark/light cycle treatments were performed as described in Chapter III.

Polymerase Chain Reaction

Total RNA was prepared from 7-day-old seedlings by phenol/chloroform extraction followed by lithium chloride precipitation as described by Lagrimini et al. (1987), and poly(A)⁺ RNA was isolated from the total RNA using a poly(A)⁺ Quick mRNA isolation kit (Stratagene, La Jolla, CA). First-strand cDNA was synthesized from 1 µg of the poly(A)⁺ RNA using Moloney Murine Leukemia Virus reverse transcriptase (Toyobo, Osaka, Japan) and was used as a template for PCR with several sets of primers (Figs.1, 2) designed from the conserved sequences between CYP71A and CYP73A2. The PCR was carried out in a reaction mixture (100 µl) consisting of 10 mM of Tris-HCl (pH 8.3), degenerate primers (at 10 µM each), 200 µM dATP, 200 µM dCTP, 200 µM dTTP, 200 µM dGTP, 2 mM MgCl₂, 50 mM KCl and 25 units/mL AmpliTaq DNA polymerase (Perkin Elmer/Cetus, Norwalk, CT). The reactions were performed through 35 cycles of 30 sec at 94°C, 30 sec at 50°C and 90 sec at 72°C using a thermal cycler (Perkin Elmer/Cetus, model 480). PCR products were separated by low-melting temperature agarose gel (1.5%) electrophoresis and were cloned into the pCRII vector using a TA cloning kit (Invitrogen). The cloned PCR products were divided into several groups by analyzing their restriction fragments from digestion with *EcoRI*, *XhoI*, and *HindIII*. The DNA sequences of the cloned PCR products were partially determined.

Isolation of cDNA clones

Each of the PCR products was labeled with [³²P]-dCTP by the random priming labeling method (Feinberg and Vogelstein, 1983). A total of 600,000 phage plaques from an Arabidopsis cDNA library (Chapter III, Mizutani et al., 1997) was screened for corresponding P450 cDNAs under high stringency conditions: hybridization, 16-h at 65°C in a hybridization buffer containing 1% BSA, 7% SDS, 50 mM sodium phosphate pH 7.5, and 1 mM EDTA (Church and Gilbert, 1984); washing, 10 min in 2 x SSC, 0.1% SDS at room temperature, and 20 min at 65°C in 0.1 x SSC plus 0.1% SDS.

Preparation and analysis of DNA and RNA

Genomic DNA and total RNA were isolated as described in Chapter III. For Southern blot analysis, the isolated full-length P450 cDNAs were used as probes, and hybridization and washing were carried out under the high stringency conditions as described in Chapter III.

RFLP mapping was performed by L.Medrano in the laboratory of E.Meyerowitz as described by Chang et al. (1988) using genomic DNA isolated from a cross between Landsberg *erecta* and Columbia ecotypes. Gel blots of genomic DNAs digested with *Xba*I were probed with the P450 cDNAs to reveal DNA polymorphisms between Landsberg and Columbia. RFLP mapping data were analyzed using the MAPMAKER computer program developed by Lander et al. (1987).

RNA gel blot analysis was performed as described in Chapter III. DNA sequencing and analysis were also carried out as described in Chapter III.

Results

PCR strategy

I isolated novel P450 clones from 7-d-old Arabidopsis seedlings by a PCR strategy. While more than 200 P450 genes from a variety of organisms had been reported (Nelson et al., 1993) when I started this study, only two plant P450s were available: avocado CYP71A1 of an unknown physiological function (Bozak et al., 1990) and a mung bean cinnamate 4-hydroxylase, CYP73A2 (Chapter II, Mizutani et al., 1993). CYP71A1 and CYP73A2 were 30% identical at the amino acid level (Fig. 1), and degenerate oligonucleotide primers were designed from five conserved regions shown in Fig. 2.

CYP71A1	MAILV---SL	L--FLAIALT	FLLKLNEKR	EKKPNLPPSP	PNLPIIGNLH	45
CYP73A2	MDLLLLLEKTL	LGLFLAAVVA	IVVSKLRGKR	<u>FK---LPPGP</u>	LPVPIFGNWL	47
MA47						
CYP71A1	QLGN-LPHRS	LRSLANELGP	LILLHLGHIP	TLIVSTAEIA	EEILKTHDLI	94
CYP73A2	QVGDDLNRHN	LTQLAKRFGD	IFLLRMGQRN	LVVVSSPDLA	KEVLHTQGV	97
CYP71A1	FASRPSTTAA	RRIFYDCTDV	AFSPYGEYWR	QVRKICVLEL	LSIKRVNSYR	144
CYP73A2	FGSRTNRNVF	DIFTGEGQDM	<u>VFTVYGEHWR</u>	KMRRIMTVPF	FTNKVVQQYR	147
MA46						
CYP71A1	SIREEEVGLM	MERISQSCST	G-EAVNLSEL	LLLLSSGTIT	RVAFGKKYEG	193
CYP73A2	HGWEAEAAAV	VDDVRKNPDA	AVSGLVIRRR	LQLMMYNMY	RIMFDRRFES	197
CYP71A1	EEERK-NKFA	DLATELTTLT	GAFFVGDY-F	PSFAVVDVLT	G-MDARLKR-	239
CYP73A2	EEDPLFQRLK	ALNGERSRLA	QSF--EYNY	GDF--IPILR	PFLKGYLKIC	242
CYP71A1	---NHGELDA	FVDHVIDDHL	LSRKANGSDG	VEQKDLVDVL	LHLQKDSSLG	286
CYP73A2	KEVKETRLKL	FKDYFVDERK	NIGSTKSTNN	EGLKCAIDHI	LDAEKKG---	289
CYP71A1	VHLNRNRLKA	VILDMFSGGT	DTTAVTLEWA	MAELIKHPDV	MEKAQQEVRR	336
CYP73A2	-EINEDNVLY	IVENINVAAI	ETTLWSIEWG	IAELVNHPEI	QQKVRDEIDR	338
CYP71A1	VVGKKAKVEE	EDLHQLHYLK	LIKETLRLH	PVAPLLVPRE	STRDVVIRGY	386
CYP73A2	VLGVGHQVTE	PDIQKLPYLQ	AVVKETLRLR	MAIPLLPHM	NLHDAKLGYY	388
CYP71A1	HIPAKTRVFI	<u>NAWAIGRDPK</u>	SWENAEFLP	ERFVNNS--V	DFKGQDFQLI	434
CYP73A2	DIPAESKILV	<u>NAWWLANNPA</u>	<u>HWKKPEEFP</u>	<u>ERFEEESHV</u>	EANGNDFRYL	438
MA40 MA48						
CYP71A1	PFGAGRRGCP	GIAFGISSVE	ISLANLLYWF	NWELPGDLTK	EDLDMSEAVG	484
CYP73A2	<u>PFGVGRRS</u> CP	GIILALPILG	ITLGRLVQNF	--ELLPPPGQ	SQIDTSEKGG	486
MA39						
CYP71A1	-ITVHMKFPL	QLVAKRHLS	502			
CYP73A2	QFSLHILKHS	TVVAKPRSF	505			

Figure 1. Comparison of the deduced amino acid sequences between CYP71A1 (Bozak et al., 1990) and CYP73A2 (Chapter II, Mizutani et al., 1993a).

Identical amino acid residues are shadowed with gray, and dashes were inserted to maximize the sequence homology. The amino acid sequences used to design degenerate oligonucleotide primers are underlined.

First, the consensus sequence for the heme-binding region (HR2 region, Gotoh and Fujii-Kuriyama, 1989) was found as PFG(A/V)GRR in CYP71A1 and CYP73A2 (Fig. 1, underlined), and a primer(MA39) was designed from this plant HR2 region (Fig. 2). Another primer (Fig. 2, MA46) was designed from a proline-rich motif (KLPPGP) found next to the N-terminal signal-anchor sequence CYP73A2 (Fig. 1, underlined). A proline-rich motif, (P/T)PGPX(P/G)XP, is known to be conserved in mammalian microsomal P450s (Yamazaki et al., 1993). Next, it was found that a short sequence, YGE(H/Y)WR, is conserved in the middle region of CYP71A1 and CYP73A2 (Fig. 1, underlined), and a forward primer was designed from this sequence (Fig. 2, MA47). A reverse primer (Fig. 2, MA48) was designed from a relatively long sequence, (A/P)EEF(R/L)PERF, conserved in the C-terminal region of

both CYP71A1 and CYP73A2. Finally, a sequence, NAWAIGRDP, from the C-terminal region of CYP71A1, was chosen to design a primer (MA40, Fig. 2). Since only 4 out of the 9 residues were consistent with the corresponding region of cinnamate 4-hydroxylase, CYP73A2 (Fig. 1, underlined), it was expected to obtain P450 clones other than CYP73. I wanted to avoid the amplification of CYP73, which is one of the most abundant P450s in plants and whose mRNA should therefore constitute a major part of plant P450 transcripts.

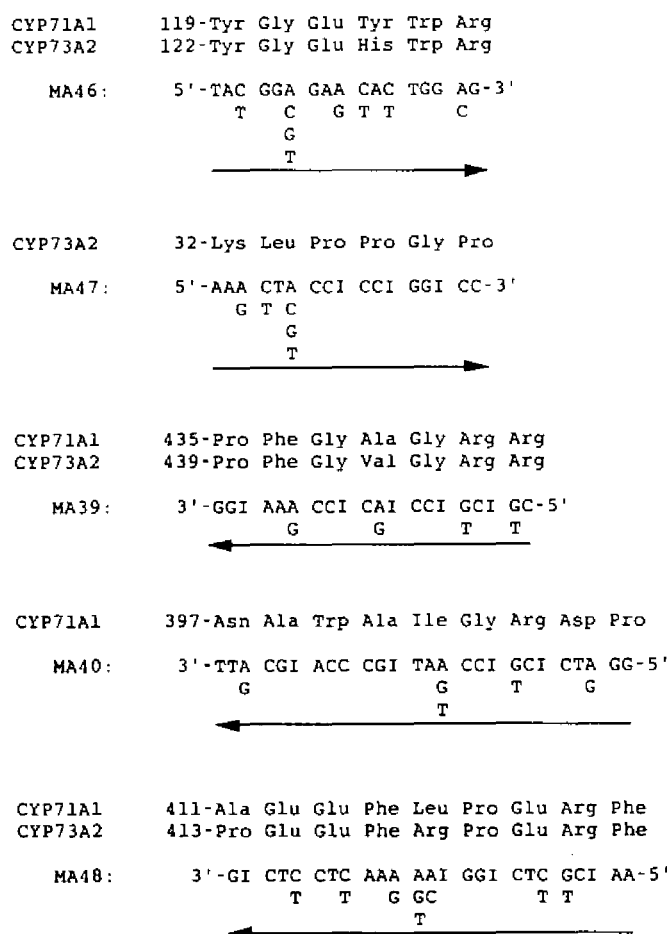


Figure 2. Degenerate oligonucleotide primers for PCR.

Degenerate oligonucleotide primers were designed from five conserved amino acid sequences between CYP71A1 and CYP73A2 (Fig. 1, underlined).

Using six primer sets (Table I), PCR was performed on first strand cDNA prepared from a poly(A⁺) RNA of 7-d-old Arabidopsis seedlings as a template, and 18 different fragments were amplified, corresponding to 11 distinct P450s (Table I). Sequence analysis revealed that 4 clones, designated P450C4H (Table I), were derived from CYP73A5, cinnamate 4-hydroxylase, from Arabidopsis (Chapter III, Mizutani et al., 1997). The

remaining 14 PCR products appeared to encode 10 novel P450s. Four classes of clones were obtained with the MA39 primer designed from the heme-binding region. On the other hand, the other 7 classes of clones were amplified using the four primers other than MA39, indicating that various homologous regions between CYP71A1 and CYP73A2 are conserved in other plant P450s.

Table I. Putative P450 cDNA fragments amplified by PCR

PCR was performed with six sets of the degenerated primers (Fig.2) using a first strand cDNA derived from a 7-d-old Arabidopsis seedlings as a template and the PCR products were cloned into a pCRII vector with a TA cloning kit.

Set of primers	Amplified P450 fragments			
MA46 x MA39	pCR-C4H-a	pCR-12-a	pCR-67-a	
MA46 x MA40	pCR-65	pCR-66	pCR-67-b	
MA46 x MA48	pCR-C4H-b	pCR-12-b	pCR-48	
MA47 x MA39	pCR-C4H-c	pCR-4-a	pCR-12-c	
MA47 x MA40	pCR-2	pCR-3	pCR-4-b	pCR-13
MA47 x MA48	pCR-C4H-d	pCR-23		

Isolation of cDNA clones

Screening of a cDNA library from 7-d-old Arabidopsis seedlings was performed using the PCR fragments as hybridization probes under high stringency conditions and obtained corresponding cDNAs (Table II). Some of the probes hybridized to multiple cDNAs: pCR-13 hybridized to a corresponding cDNA (P450-13-6) and three additional different cDNAs (P450-13-1, 13-5 and 13-7); pCR-66 hybridized to its corresponding cDNA (P450-66-8) and an additional one, P450-66-5; pCR-67 hybridized to its corresponding P450-67-3 and another cDNA clone, P450-67-1. The remaining PCR clones hybridized only to their corresponding cDNAs. DNA sequences of 11 of the 16 cDNAs were completely determined, and the deduced primary structures are shown in Fig. 3. These P450s were assigned to five distinct gene families (see below and Table II). Sequence comparison (data not shown) indicated that P450-65 cDNA lacked a few bases at the 5'-terminus of a corresponding full-length cDNA. Partial DNA sequencing showed that the remaining 5 clones, P450-66-5, P450-67-1, and P450-67-3 also represent full-length cDNAs and P450-

13-5 and P450-13-7 are partial clones (data not shown). Thus, I have isolated 13 full-length and 3 partial cDNAs encoding 16 distinct P450s from *Arabidopsis*.

Table II. P450 cDNA clones isolated using the corresponding PCR clones as probes

PCR clones	P450 cDNA clones	Family
pCR-C4H-a	P450-C4H	CYP73A5
pCR-2	P450-2	CYP83B1
pCR-3	P450-3	CYP83A1
pCR-4-a	P450-4	CYP76C1
pCR-12-a	P450-12	CYP71B5
pCR-13	P450-13-1	CYP71B3
	P450-13-5	ND
	P450-13-6	CYP71B4
	P450-13-7	ND
pCR-23	P450-23	CYP71B6
pCR-48	P450-48	CYP71B2
pCR-65	P450-65	CYP91A1
pCR-66	P450-66-5	ND
	P450-66-8	CYP91A2
pCR-67-a	P450-67-1	ND
	P450-67-3	ND

ND: Not determined

Multiple P450 gene families in Arabidopsis

1. CYP71

Five P450 cDNAs (P450-12, 13-1, 13-6, 23, and 48) were 44-69% identical at the amino acid level to one another (Table III), and their deduced amino acid sequences were more than 40% identical to that of avocado CYP71A1. These P450s were assigned to the CYP71B subfamily as follows: P450-12, CYP71B5; P450-13-1, CYP71B3; P450-13-6, CYP71B4; P450-23, CYP71B6; P450-48, CYP71B2. CYP71B3 and CYP71B4, isolated using the same probe (pCR-13), showed the highest identity to each other at both the amino acid (68.6%) and the DNA levels (76.7%) each other. Twenty-two P450s belonging to the CYP71 family have so far been isolated from several plant species (Nelson et al., 1996). However, none of their physiological functions have yet been identified. A high degree of diversity of the predicted protein sequences of the *Arabidopsis* CYP71B P450s (Fig. 3) suggests distinct catalytic activities.

2. CYP73

The P450-C4H cDNA (Table II) was completely identical to CYP73A5-encoding cDNA (cinnamate 4-hydroxylase) previously isolated from *Arabidopsis* (Chapter III, Mizutani et al., 1997). P450 clones belonging to CYP73 have also been isolated from several plant species: CYP73A1 from *Helianthus tuberosus* (Tuetsch et al., 1993), CYP73A2 from *Phaseolus aureus* (Mizutani et al., 1993), CYP73A3 from *Medicago sativa* (Fahrendorf and Dixon, 1993), CYP73A4 from *Catharanthus roseus* (Hotze et al., 1995), CYP73A9 from *Pisum sativum* (Frank et al., 1996), and CYP73A10 from *Petroselinum crispum* (Logemann et al., 1995). It has already been shown from heterologous expression experiments that 5 P450s belonging to the CYP73 family (CYP73A1 through CYP73A5) have cinnamate 4-hydroxylase activity (Chapter III, Tuetsch et al., 1993; Fahrendorf and Dixon, 1993; Hotze et al., 1995, Mizutani et al., 1997). While the CYP73 P450s so far isolated are 75-90% similar to one another at the amino acid level, it is not clear whether or not all the members of CYP73 catalyze the same reaction. It has been reported that single amino acid substitutions in P450 proteins can result in dramatic changes of substrate specificity (Lindberg and Negishi, 1989; Johnson, 1992).

3. CYP76

The deduced primary structure of P450-4 was 26 to 37% homologous to those of the other *Arabidopsis* P450s (Table III). The highest identity (40%) was found with CYP76A2, isolated from *Solanum melongena* hypocotyls using a CYP75 cDNA as a probe (Toguri et al., 1993). P450-4 was therefore designated to a new P450 subfamily, CYP76C1. Four P450s in this family have been isolated (Nelson et al., 1996), but their physiological functions are still unknown.

4. CYP83 and CYP91

Sequence comparison showed that P450-3 was completely identical to CYP83A1 previously isolated by cross-hybridization with a gene rescued from a T-DNA-tagged *fah1* mutant of *Arabidopsis*, which lacks ferulate 5-hydroxylase activity (Chapple 1995; Meyer et al., 1996). The amino acid sequence deduced from the P450-2 cDNA was 61.6% identical to

that predicted from the P450-3 cDNA, and P450-2 was assigned to a new subfamily, CYP83B1.

P450-65 and P450-66-8, which are 50.4% identical at the amino acid level (Table III), have been assigned to a new P450 family, CYP91A1 and CYP91A2, respectively. No P450s belonging to this family have been reported from other plant species.

Table III. Homology of DNA and amino acid sequences of the P450s isolated from Arabidopsis

	71B2	71B3	71B4	71B5	71B6	73A5	76C1	83A1	83B1	91A1	91A2
	Nucleotide identity (%)										
71B2 (P450-48)	-	65.8	66.2	63.7	55.0	49.2	50.6	52.1	54.1	45.6	48.1
71B3 (P450-13-1)	57.3	-	76.1	60.8	50.6	49.5	49.9	48.7	49.8	47.2	48.8
71B4 (P450-13-6)	57.8	68.6	-	61.9	56.5	49.8	50.7	53.4	53.1	48.0	47.9
71B5 (P450-12)	54.1	54.1	51.0	-	54.6	49.6	48.9	51.0	52.9	44.1	46.0
71B6 (P450-23)	43.6	45.0	46.1	46.8	-	49.0	50.1	50.9	52.8	48.0	47.9
73A5 (P450-C4H)	26.2	28.9	32.0	29.3	27.9	-	49.0	50.4	49.8	46.2	47.8
76C1 (P450-4)	34.6	34.9	35.4	31.5	36.2	25.8	-	50.4	51.0	47.1	48.5
83A1 (P450-3)	36.9	37.3	37.7	30.8	36.0	29.9	33.0	-	63.5	46.6	50.0
83B1 (P450-2)	40.2	41.2	39.3	36.9	40.2	28.5	34.4	61.6	-	48.6	48.4
91A1 (P450-65)	33.5	33.3	34.8	32.1	31.9	28.2	32.0	29.7	30.0	-	57.6
91A2 (P450-66-8)	31.9	32.8	32.9	31.0	30.0	25.3	32.0	29.6	32.1	50.4	-
	Amino acid identity* (%)										

* Amino acid identities higher than 40% are shown in bold type.

Of 74 P450 families from both eukaryotes and prokaryotes, 23 families (CYP51 and CYP71 through CYP92) are represented in higher plants (Nelson et al., 1996). The P450s isolated in this study belong to either established subfamilies; CYP71B, CYP73A and CYP83A, or novel subfamilies; CYP76C, CYP83B and CYP91A. Thus, a total of 10 families, including those so far reported (CYP72, CYP74, CYP84, CYP86 and CYP90) have been now established in Arabidopsis.

5. Protein structures

The P450s isolated in this study have significant high homology throughout the overall sequences including 35 conserved amino acid residues (Fig. 3). In addition, each P450 contains a hydrophobic stretch at its N-terminus (Fig. 3), which has the structural properties of the signal-anchor sequence of microsomal P450s (Nelson and Strobel, 1988). The region contains a few charged residues and a proline rich motif following the hydrophobic stretch, both of which are responsible for the targeting and anchoring of newly

synthesized P450 proteins at the cytoplasmic surface of the endoplasmic reticulum (Yamazaki et al., 1993; Beltzer et al., 1991). An extended consensus sequence of the core heme-binding site, FxxGxR/HxCxG, was identified as P/AFGxGRR/KxC^{PG}/A in these Arabidopsis P450s (Fig. 3, underlined). Interestingly, the highly conserved glycine found 2 residues C-terminal to the heme-ligating cysteine was substituted for alanine in CYP71B6 and CYP83B1. Of more than 400 P450 sequences so far determined, this substitution has been found in only 10 P450s from insect, fungi, and plants (Nelson et al., 1996). It is not known how the substitution of this structurally important glycine influences the active site conformation.

Genomic Southern blot analysis

I have previously reported that CYP73A5 is encoded by a single copy gene in Arabidopsis (Chapter III, Mizutani et al., 1997). To estimate the copy number of the other 10 P450 genes described here, I performed genomic Southern blot analyses under the high stringency conditions (65 °C in 0.1 x SSC and 0.1% SDS). I observed distinct hybridization patterns unique to each P450 gene (Fig. 4), which indicated that the individual hybridization probes did not cross-hybridize to other P450 genes. The hybridization patterns in the genomic Southern analysis (Fig. 4) were consistent with those predicted from the restriction patterns of the corresponding cDNAs (data not shown), suggesting that each P450 is encoded by a distinct single copy gene in Arabidopsis.

CYP71B2	MTI-----	LLCFFLVSL	LTIVSS-IFL	KONKTSK--F	NLPPSPSS-L	PIIGNLHHLA	GL-PHRCFKH	LSIKY-----	GPLVFLRLGS	VPVVVISSE	82	
CYP71B3	MSI-----	LLYFFFL--	-PVILSLIFM	KKFKDSK--R	NLPPSPPK-L	PIIGNLHQLR	GL-FHRCLEHD	LSKKH-----	GPVLLRLRGF	IDMVVISSE	80	
CYP71B4	MVS-----	LLSFFLLLL	VPFFLLIFT	KKIKESK--Q	NLPPGPKG-L	PIIGNLHQLR	GL-LHKCLHD	LSKKH-----	GPVHLRLRGF	APMVVISSE	83	
CYP71B5	MSI-----	FLCPLLLLP	LS----LIFL	KKLLPSK--G	KLPPGPKG-L	PIIGNLHQLR	RF-LHKSLLK	ISOEY-----	GPVHLHFGV	VPVVISSE	79	
CYP71B6	MSLFFSPIST	ELLFWLLLL	IPPLIFLFL	RSPK-----	NLPPGPPR-L	PIIGNHOGF	SL-PHRSRLD	LSLKY-----	GPVITVYGLS	VRIVVHSP	87	
CYP73A5	NDLILLLEKS	-----LIIV	FVAVILATVI	SKLRGKK--L	KLPPGPI-P-I	PIFGNWLQVG	DDLNRHLVD	YAKKF-----	GDLLFLRMGO	RNLVVSPPD	85	
CYP76C1	MDI-----	ISG	QALLLLFCFI	LSCFLITFTT	RSGRISRGAT	ALPPGPPR-L	PIIGNIHLV	GKHPHRSFAE	LSKTY-----	GPVMSLKLGS	LNTVVIASPE	89
CYP83A1	MEDIITGVVA	-----	LAAVLLPFLY	QKPKTKR--Y	KLPPGSP-I-L	PIIGNLHQLR	KFNPHOFLFR	WAKKY-----	GPILSYRIGS	RTMVVISAE	82	
CYP83B1	MDLLLI-TAG	-----	LVAAAAFFFL	RSTTKAS--L	RLPPGPKG-L	PIIGNLHQLR	KFNPHOFLFR	LSKLY-----	GPITFMKIGG	RRIAVISSE	81	
CYP91A1	ETNI-----	RVLVYSIFSL	IFLIISFKFL	KPKKO-----	NLPPSPGWL	PIIGHLR-LL	KPIIHRLLRS	FSETLDHNDG	GGVMSLRLGS	RLVYVSSHK	88	
CYP91B1	M-----	LYFILLPL	LFLVISYKFL	YSKTQR---F	NLPPGPPSR-	PFVGHLL-LH	KPIIHRLLOR	YSNOY-----	GPISLRFSGS	RRVVVITSPS	79	

CYP71B2	AA-EAVLKTN	DLECCSRPKT	VSGSKLSYGF	KDIT-FAPYG	EYWEVRKLA	VIELFSSKKV	QSPRYIREEE	VDFVVKVSE	SALKQ--SP-	---VDLSKTF	174	
CYP71B3	TT-EVLKXVH	DLECCRPKT	NASSKFSRDG	KDIA-FAPYG	EVSRELRKLS	LINFSTOKV	RSPRYIREEE	NOLMVKKLKE	SAKKK--NT-	---VDLSOTL	172	
CYP71B4	AA-EALKTH	DLECCSRPT	MASRVFSRNG	KDIG-FGYVG	DENRELRKLS	VREFPSVKKV	QSPRYIREEE	NOLMVKKLKE	LASKO--SP-	---VDLSKIL	175	
CYP71B5	GA-EVLKTH	DLECCSRPKT	VSGSLFTYFN	KDIG-FAPYG	ENWREMRKIA	VSELPSQKLL	KSPRYIREEE	SQLLVKRVSK	SALEPTSS-	---VNLKRII	173	
CYP71B6	TA-EVLKLR	DSECCTRPKL	SITKSFFYDG	LGLG-FTXWG	DYRDVVRKLC	VLELPSVKRA	NSFRNIREEE	LSRLVNSFSD	SA--SSGSS	---VDLTANI	179	
CYP73A5	LT-KEVLQTO	GVEFGSRTRN	VYFDIPTGKG	QDMV-PTYVG	ENWREMRKIM	TVPFSTKRVV	QONREGWEFE	AASVVEDVKK	NPSDATKG-	---IVLKRRL	178	
CYP76C1	AA-REVLRTH	DQILSARSPT	NAVRSINHOD	ASLVWLPSSS	ARMRLRLRLS	VTOLLSPORI	EATKALRMNK	VKELVSFISE	SSDRE--ES-	---VDISRAV	182	
CYP83A1	LA-KELLKTO	DVNFADRPPT	RAHEPISYGR	RDMA-LNHYT	PYREIRKMG	MNHLFSPTRV	ATFKHVRREE	ARRMMDKINK	AADKS--EV-	---VDISELM	174	
CYP83B1	LA-KELLKTO	DVNFADRPPT	RAHEPISYGR	RDMA-LNHYT	PYREIRKMG	MNHLFSPTRV	ATFKHVRREE	ARRMMDKINK	AADKS--EV-	---VDISELM	174	
CYP91A1	VAAEECFGKN	DVVLANRPQV	IIGKHVGYNN	ANMI-AAPYG	DHWRNLRRLC	TVLFSFTHRL	NCPLYVRTDE	VRRL-SRLSR	LAGTKK----	---TVVLEKPL	182	
CYP91B1	LAQESFTGQN	DIVLSRRPLQ	LTAKYVVYNN	TTVG-TAPYG	DHWRNLRRLC	SQETLSSHRL	IIFQHTRLRS	ILRMLTRLRS	YTQTSNESND	FTHIELEPLL	178	

CYP71B2	FSLTASIIICR	VALGONF--N	ESGFVIOQDR	IEELVTESAE	ALGTFTPSDF	FP-GGLGRFV	DWLFORHKKI	NKVFKELOAF	YQHVIDDBLK	---PEGRKNO	268	
CYP71B3	FYLVGSIIFR	ATFGQRL--D	QNKHV-NKEK	IEELMFEV-O	KVGSLSSSDI	FP-AGVGMFM	DFVSGRHKTL	HKVFEVDTL	LNVHIDGRLK	---NPEDXTN	264	
CYP71B4	FGLTASIIIFR	TAFGQSF--F	DNKHV-DOES	IKELMFES-L	SNMTFRPSDF	PPTAGLKKWFI	GFVSGQHKRL	YVNVFVDTF	FNVHIDDBLK	---HSSKAT	265	
CYP71B5	FTFAASIIICR	LSFGQNF--C	D--FVDMET	VEELVLES-ET	NLGSAPADP	LPAC--HII	DRISGOSTV	HKAFSKLTFN	FELVIDDBLK	---SGK--I	259	
CYP71B6	AKFVASFTCR	MAFGLSF--O	GSG--MDNET	FLELFTANR	VIGKPAADI	FP--GFGWIL	DRISGLDSSR	RKSPQDLDT	YQVATVDR--	---REKKKT	267	
CYP73A5	QLMYTIHNR	IMFDRRPESE	DDPLFLRLKA	LNGERSRLAQ	SFE-YNYGDF	IP1--LRPF-	--LRGLKIC	ODVKKRRIAL	FKKYFVDERK	QIASSKPTGS	272	
CYP76C1	FITLNIISN	ILFRVQL--G	SYNAKASING	QDVTIVSMQ	AAGTPDAANY	FPFL--RFL	D--LQGNVKT	KVCTERLVRV	FRGFIIDAKIA	EKSSQNNPKD	262	
CYP83A1	LTFTNSVCCR	QAFGKKY--N	EDGE--EMKR	FIDILYGTQS	VLGKIFPSDF	FPYCG--FL	DDLGLTAYM	KECFERODTY	IOEVVNETLD	---P--KRVK	261	
CYP83B1	LSFTNCVCCR	QAFGKKY--N	EYGT--EMKR	FIDILYGTQS	VLGKIFPSDF	FPYCG--FL	DDLGLTAYM	KECFERODTY	IOEVVNETLD	---P--KRVK	261	
CYP91A1	MDLTENNIMR	MHTGKRYTG	ETDDEEAKR	VRKLVDVGA	NTSSGNAVYD	VPI--LRLES	S-Y--ENRV	KKLGEETDIF	LOGLIDD---	---KRGQO	268	
CYP91B1	SDLTENNIVR	MVTGKRYTG	DVNNKEEAE	PKKLVDYDAM	YSQANKSADY	LPI--LKLPG	NKF--EKEV	KAIGKSMDDI	LQRLDDE---	---CRDRK	265	

CYP71B2	---DIVTLI	LDMDIKOEDS	DSFKLNMDNL	KAVMDVFLA	GIOTSAVTMI	WAMTELIRNP	RVMKKAQGS	RTTLGLK-KE	RITVEDLGVK	EYLNHILKET	363	
CYP71B3	QDRPDIDSI	LETIYQKQD	ESFKLTIDHL	KGIQNIYLA	GVDTSAITMI	WAMTELIRNP	RVMKKAQGS	RTTLGLK-KE	RITVEDLGVK	EYLNHILKET	364	
CYP71B4	QDRPDIDSI	LETIYQKQD	ESFKLTIDHL	KGIQNIYLA	GVDTSAITMI	WAMTELIRNP	RVMKKAQGS	RTTLGLK-KE	RITVEDLGVK	EYLNHILKET	364	
CYP71B5	EDHSDIISVM	LDMDIKOEDS	DSFKLNMDNL	KAVMDVFLA	GIOTSAVTMI	WAMTELIRNP	RVMKKAQGS	RTTLGLK-KE	RITVEDLGVK	EYLNHILKET	363	
CYP71B6	EDREDIDIVL	LKLGSOETKL	GSSRITDTHI	RAIMDLVPA	GVDTSVITLD	WAMTELIRNP	RVMKKAQGS	RTTLGLK-KE	RITVEDLGVK	EYLNHILKET	363	
CYP73A5	EGELKCAIDHI	LEAQKQGE--	---INEDNV	LYIVENINVA	AIETTLMSIE	WGIATLVNHP	EIOSKLGNEH	DTVLPGCVQ-	-VTEPLHKL	PYLQAVVKET	364	
CYP76C1	VSKNDPVDNL	LD--YKGDSE	E--LSISDI	EHLLDMFTA	GTDTSSSTLE	WMTTELKLRP	KTKAKAQAEI	DCVIG--QNG	IVEESDISKL	PYLQAVVKET	369	
CYP83A1	PETESMIDLL	MG-YKQDPF	-ASEFTVDNR	KAVILDIIVA	GTDTAAAVV	WMTTELKLRP	KTKAKAQAEI	DCVIG--QNG	IVEESDISKL	PYLQAVVKET	369	
CYP83B1	OETESMIDLL	MG-YKQDPF	-ASEFTVDNR	KAVILDIIVA	GTDTAAAVV	WMTTELKLRP	KTKAKAQAEI	DCVIG--QNG	IVEESDISKL	PYLQAVVKET	369	
CYP91A1	ETGTIMIDHI	LVLQKSDIEY	---YTDQII	KGIILINVA	GTNTSAVTE	WAMTELIRNP	RVMKKAQGS	RTTLGLK-KE	RITVEDLGVK	EYLNHILKET	363	
CYP91B1	E-GNTMVNHL	ISLQOOQPEY	---YTDVII	KGLHMSHMLA	GTETSATVLE	WAMTELIRNP	RVMKKAQGS	RTTLGLK-KE	RITVEDLGVK	EYLNHILKET	363	

CYP71B2	FRLHPALPFV	VPRETMADIK	IGQYDIPPKT	OIGLVNMTIG	RDPKRW-NDP	EEFNPERFAN	SS--VDFRGO	HFDLLPFGSG	RRICPGMPMA	IASVELALMN	460	
CYP71B3	LRLHPPAPLL	LPRETMADIK	IGQYDIPRKT	ILLVNAWSIG	RNPFLW-ENP	EEFNPERFID	CP--MDYKGN	SFEMLPFGSG	RKICPGIATG	IATVELGLLN	461	
CYP71B4	LRLHPPAPLL	LPRETMADIK	IGQYDIPRKT	ILLVNAWSIG	RNPFLW-ENP	EEFNPERFID	CP--MDYKGN	SFEMLPFGSG	RKICPGIATG	IATVELGLLN	461	
CYP71B5	FRLHPPAPLL	LPRLTMSDIN	IQGYDIPKNT	HIQINTYITG	RDPKRW-TKP	DEFIPERFVD	NP--IYKQGO	HFELLPPFAG	RRVCPGMATG	ITIVELGLLS	455	
CYP71B6	WRLHAPSPIL	IPREANTNFK	IKGYDIYPTG	RHVNNAWIA	RNPKNW-KDP	DEFIPERFVD	SN--VETKGT	SFELLPPFAG	RRVCPGMATG	ITIVELGLLS	455	
CYP73A5	WRLHAPSPIL	IPREANTNFK	IKGYDIYPTG	RHVNNAWIA	RNPKNW-KDP	DEFIPERFVD	SN--VETKGT	SFELLPPFAG	RRVCPGMATG	ITIVELGLLS	455	
CYP76C1	FRLHPPAPLL	IPRKAESDAE	ILGFMVLDKT	QVLVNVMAIG	RDPKRW-NDP	DEFIPERFID	NP--IYKQGO	HFELLPPFAG	RRVCPGMATG	ITIVELGLLS	455	
CYP83A1	LRIEPIVPIPL	IPRACIQDITK	IAGYDIPAGT	TVNVNAWAVS	RDEKEMGPNP	DEFIPERFID	NP--IYKQGO	HFELLPPFAG	RRVCPGMATG	ITIVELGLLS	455	
CYP83B1	LRIEPIVPIPL	IPRACIQDITK	IAGYDIPAGT	TVNVNAWAVS	RDEKEMGPNP	DEFIPERFID	NP--IYKQGO	HFELLPPFAG	RRVCPGMATG	ITIVELGLLS	455	
CYP91A1	LRLRPATPILL	VPNHASEDCR	IGSYDMPRGT	TLVNAWIAH	RDPNTW-DDP	DSFKPERF-E	KEEBAO----	---KLAPGLG	RRACPGSGLA	QRTVGLALGS	454	
CYP91B1	FRLFPVAPFL	IPRSPDDMK	IGGYDIPRDT	IVNVNAWIAH	RDPNTW-DDP	DSFKPERF-E	KEEBAO----	---KLAPGLG	RRACPGSGLA	QRTVGLALGS	454	

CYP71B2	LLYFYDWSMP	DGTGKEDIDM	EEAGNISIVK	KIPLQLV---	-PV--QRY--	502						
CYP71B3	LLYFYDWSMP	DEDK--DIDM	EEAGDATIVK	KVPLELV---	-PI--IH-H-	501						
CYP71B4	LLYFYDWSMP	EKKK--DMDM	EEAGDLTVDK	KVPLELL---	-PV--IRISL	504						
CYP71B5	LLYFYDWSMP	NGMTTKDIDM	EEAGAFVIAK	KVSELEV---	-PT--LR-RW	498						
CYP71B6	LLYFYDWSMP	---KATEEVS	EEAGGLTSHR	KHPLHLV---	-PVNVINRKL	503						
CYP73A5	MYONFEL-LP	PPGQSKVDS	EKGGOFSLRI	LNNHSIIVMKP	RNC-----	505						
CYP76C1	LLYFYDWSMP	KGVLSL---	---GFGHOR	DLWSNFA---	-----	495						
CYP83A1	LLLSFNFKLP	NGMKPDDINM	DVMTGLAMHK	SKRLKLV---	-PEKVVNYK--	502						
CYP83B1	LLYFYDWSMP	KGKIPEDIKM	DVMTGLAMHK	SKRLKLV---	-PTK--HI--	499						
CYP91A1	LQOCFEW---	ERVGNVEVDM	KEGVGNTPVK	AIPLKAICKA	RPLFKHII-S	500						
CYP91B1	LQOCFEW---	ENVKGEEMDM	SESTGLGMRK	MDPLRAMCRP	RPIMSKILL-	500						

Figure 3. Amino acid sequence alignment for the 11 P450 cDNAs from Arabidopsis. The amino acid sequences deduced from the 11 P450 cDNA sequences were aligned by the program, Clustal W (Thompson et al., 1994). Identical amino acid residues are shadowed with a gray color, and dashes are inserted to maximize the sequence homology. Proline rich motif and the heme-binding domain (HR2 region, Gotoh and Fujii-Kuriyama, 1989) are underlined. Amino acid residues conserved among the all P450s are indicated by asterisks under the alignment.

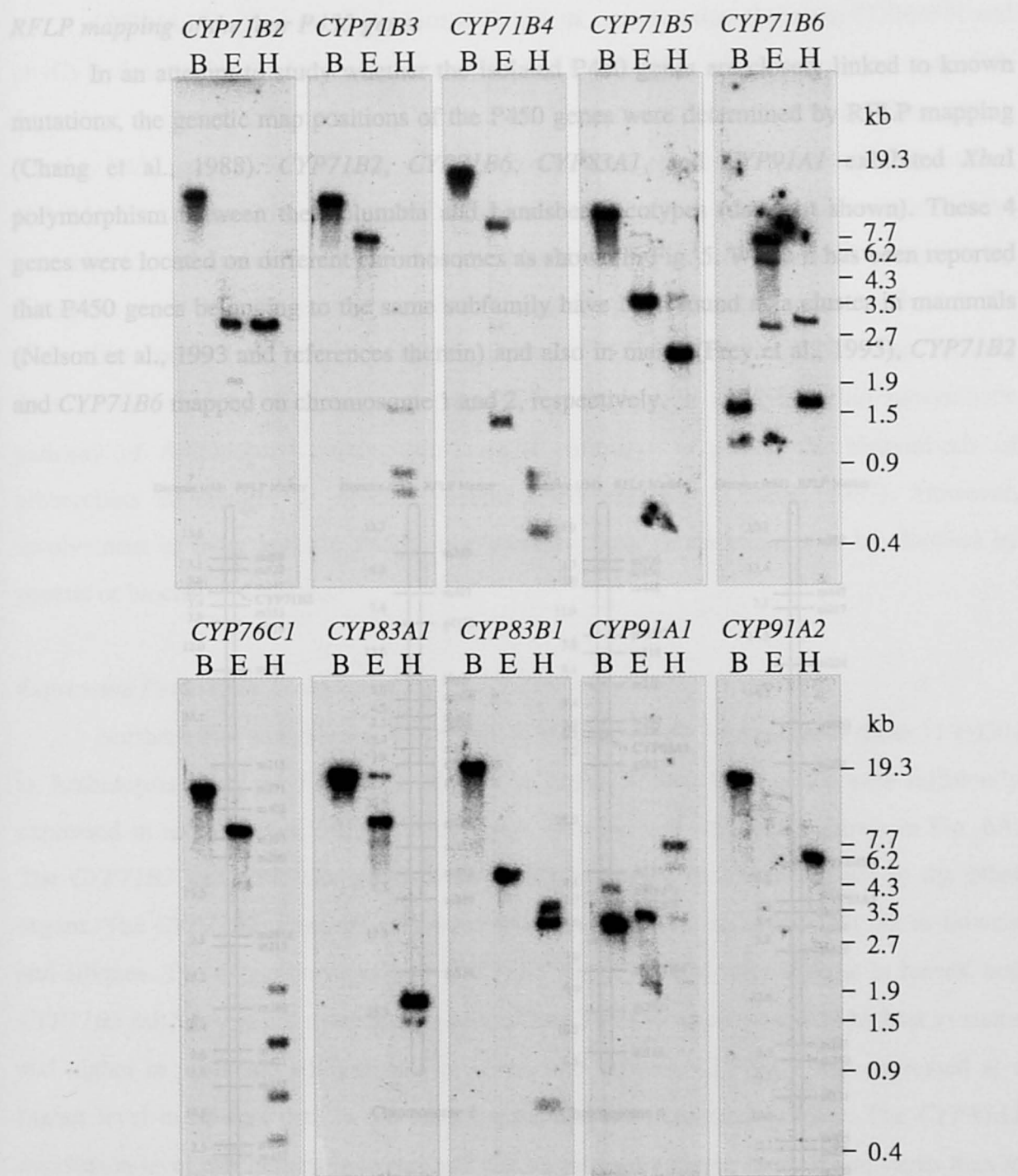


Figure 4. Genomic Southern blot analysis of 10 P450 genes of *Arabidopsis*. One microgram of *Arabidopsis* Columbia genomic DNA was digested with the restriction enzymes; B: *Bam*HI, E: *Eco*RI, and H: *Hind*III. The digested genomic DNA was separated on 0.7% agarose gel, blotted on a nylon membrane, and hybridized with [³²P]-labeled full-length cDNAs of the 10 P450s indicated. The migration of size markers is shown to the right of the blots.

RFLP mapping of the four P450 genes

In an attempt to study whether the isolated P450 genes are closely linked to known mutations, the genetic map positions of the P450 genes were determined by RFLP mapping (Chang et al., 1988). *CYP71B2*, *CYP71B6*, *CYP83A1*, and *CYP91A1* exhibited *Xba*I polymorphism between the Columbia and Landsberg ecotypes (data not shown). These 4 genes were located on different chromosomes as shown in Fig. 5. While it has been reported that P450 genes belonging to the same subfamily have been found as a cluster in mammals (Nelson et al., 1993 and references therein) and also in maize (Frey et al., 1995), *CYP71B2* and *CYP71B6* mapped on chromosome 1 and 2, respectively.

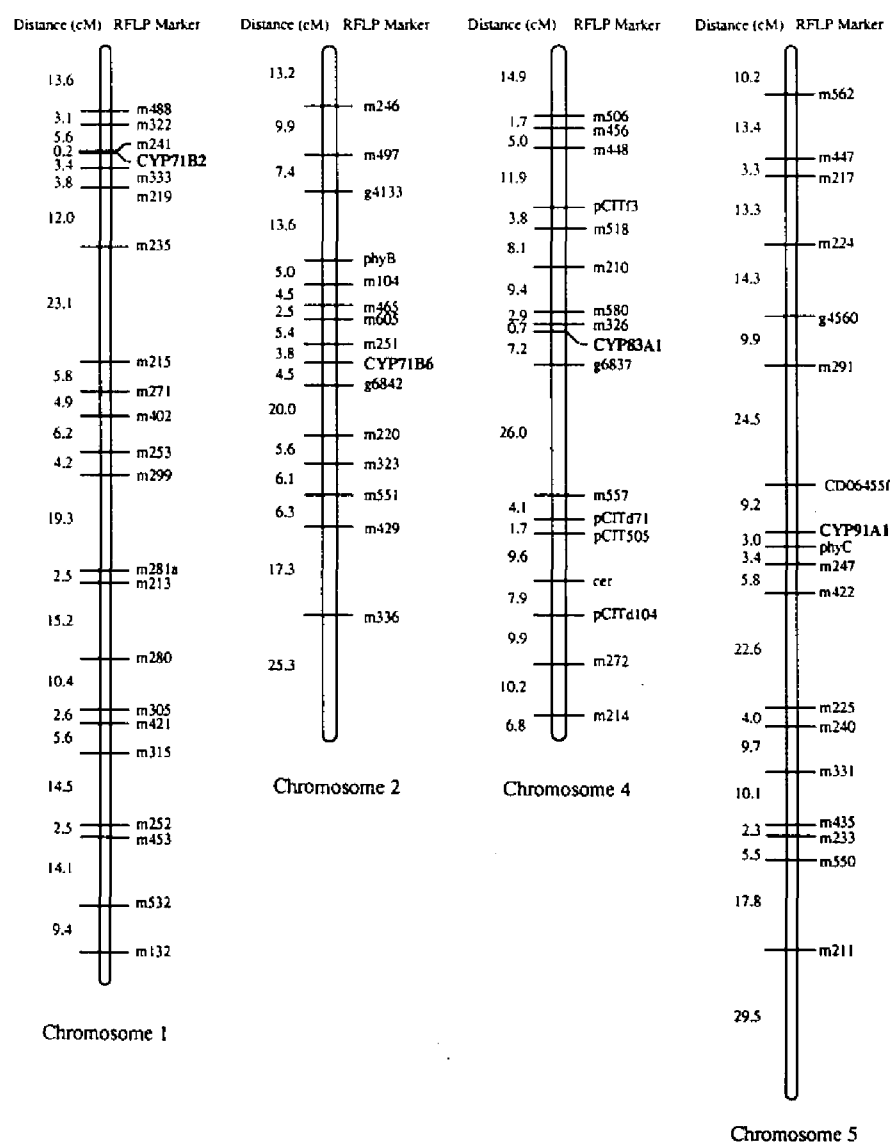


Figure 5. RFLP mapping of the *CYP71B2*, *CYP71B6*, *CYP83A1*, and *CYP91A1* genes. RFLP markers and the four P450 genes are shown on the schematic Arabidopsis genetic map. The numbers at left indicate the map distance (in centimorgans) between markers.

CYP91A1 mapped on chromosome 5, and its map position (between CD06455f and phyC) is close to the *ga3* locus (between m291 and CD06455f), which is involved in a P450-dependent oxidation of kaurene in gibberellin biosynthesis (Zeevaart and Talon, 1991). Another P450 gene, *CYP71B2* mapped at the position 0.2 cM downstream an RFLP marker, 1bAt241, on chromosome 1. It should be noted that the *ga4*, *gai*, and *dwl* mutations have been also located appropriately 0.6 cM, 0.8 cM and 1.2 cM distant from m241, respectively (Hauge et al., 1993). These results are implicating the possibility that genes involved in the gibberellin metabolism might exist as a cluster. Chiang et al. (1995) has recently demonstrated that the *GA4* gene encodes 3 β -hydroxylase in the gibberellin biosynthetic pathway of Arabidopsis. Among biochemical pathways in plants, the biosynthesis of gibberellins is thought to involve several P450 enzymes (Graebe, 1987). However, involvement of these specific P450s in gibberellin metabolism remains to be clarified by genetic or biochemical means.

Expression Patterns in Arabidopsis

Northern blot analysis was performed to investigate the expression of these 11 P450s in Arabidopsis using the full-length cDNAs as probes. These P450 genes were differently expressed in roots, leaves, inflorescence stems, flowers, and siliques as shown in Fig. 6A. The *CYP71B3* and *CYP71B4* genes were highly expressed in leaves but not in the other organs. The *CYP71B2* transcript was detected in roots, leaves, and stems but not in flowers and siliques. The expression levels of *CYP71B5* and *CYP71B6* were highest in leaves, and *CYP71B5* mRNA was not detectable in roots. The *CYP73A5* expression was highest in stems and higher in roots and siliques than in leaves and flowers. *CYP76C1* was expressed at a higher level in flowers than in the other organs and not expressed in roots. The *CYP83A1* expression level was highest in leaves and was significantly higher in roots and stems than in flowers and siliques. In contrast to relatively low expression levels of the other P450 genes in roots, the *CYP83B1* gene was most strongly expressed in roots. The *CYP91A1* mRNA level was highest in leaves compared with that in the other organs, while *CYP91A2* mRNA was strongly expressed in flowers. The results demonstrated that each of the P450 genes, even within the same P450 gene subfamilies, showed unique expression patterns.

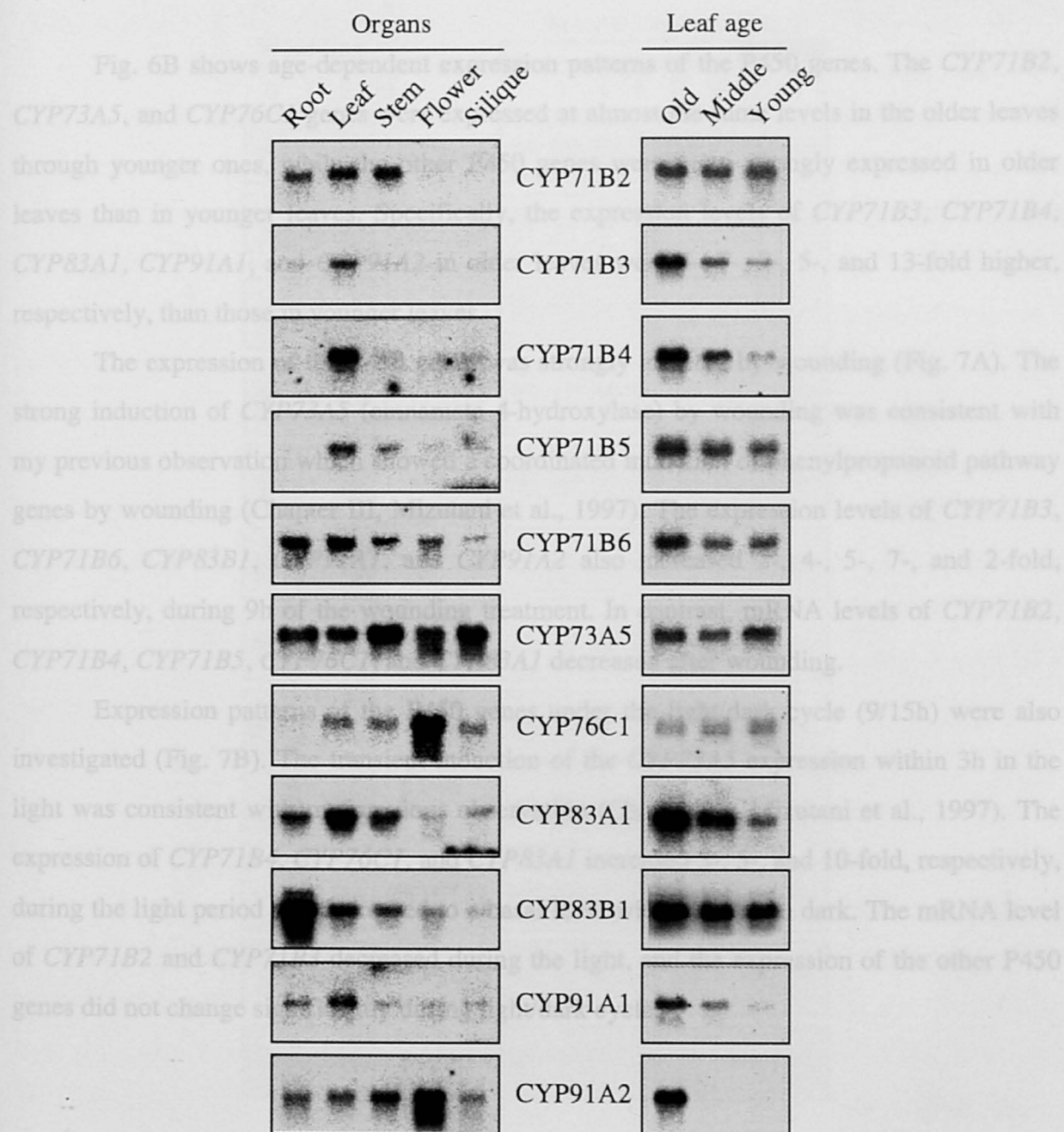


Figure 6. Tissue specific expression of the 11 P450 genes in Arabidopsis.

(A) Total RNA was isolated from the roots and leaves of 3-week-old plants, from inflorescence stems and flowers of 4-week-old plants, and from the siliques of 5-week-old plants.

(B) Total RNA was isolated from the leaves of different age of 3-week-old plants with 12 leaves. The first and second leaves represented the older leaves. The middle-aged leaves were collected from the fourth and fifth positions, and the younger leaves were from the ninth and tenth positions counted from the bottom. Plants were grown under continuous light. Five μ g of total RNA was separated on formaldehyde agarose gels, transferred to nylon membranes, and hybridized to the indicated probes.

Fig. 6B shows age-dependent expression patterns of the P450 genes. The *CYP71B2*, *CYP73A5*, and *CYP76C1* genes were expressed at almost the same levels in the older leaves through younger ones, while the other P450 genes were more strongly expressed in older leaves than in younger leaves. Specifically, the expression levels of *CYP71B3*, *CYP71B4*, *CYP83A1*, *CYP91A1*, and *CYP91A2* in older leaves were 4-, 7-, 9-, 5-, and 13-fold higher, respectively, than those in younger leaves.

The expression of the P450 genes was strongly affected by wounding (Fig. 7A). The strong induction of *CYP73A5* (cinnamate 4-hydroxylase) by wounding was consistent with my previous observation which showed a coordinated induction of phenylpropanoid pathway genes by wounding (Chapter III, Mizutani et al., 1997). The expression levels of *CYP71B3*, *CYP71B6*, *CYP83B1*, *CYP91A1*, and *CYP91A2* also increased 2-, 4-, 5-, 7-, and 2-fold, respectively, during 9h of the wounding treatment. In contrast, mRNA levels of *CYP71B2*, *CYP71B4*, *CYP71B5*, *CYP76C1*, and *CYP83A1* decreased after wounding.

Expression patterns of the P450 genes under the light/dark cycle (9/15h) were also investigated (Fig. 7B). The transient induction of the *CYP73A5* expression within 3h in the light was consistent with my previous observation (Chapter III, Mizutani et al., 1997). The expression of *CYP71B4*, *CYP76C1*, and *CYP83A1* increased 3-, 5-, and 10-fold, respectively, during the light period and decreased to a basal level within 3h in the dark. The mRNA level of *CYP71B2* and *CYP71B3* decreased during the light, and the expression of the other P450 genes did not change significantly during light/dark cycle.

Discussion

I have isolated and identified 11 P450 clones from Arabidopsis. These include 10 cytochrome P450 clones and one actin gene. The 10 cytochrome P450 clones have been grouped into 11 families based on their sequence homology and several of these families have been further characterized. Arabidopsis in this study was grown under continuous light and screened only a small number of clones. It is possible that many P450 clones in Arabidopsis are not expressed under high stringency conditions. A libraries prepared from either different tissues or under different conditions (wounding, pathogen infection, or UV irradiation) may reveal more P450 genes. These 10 found in Arabidopsis have been compared with the number of cloned P450 genes will increase through the Arabidopsis EST data base searches. The physiological functions of the identified P450 genes will be determined by the efforts of other researchers. Expression of the 11 P450 genes have physiological implications. For example, CYP73A5 (CYP73A5) is induced by light with the gene CYP73A5 (CYP73A5) pathway such as phenylalanine ammonia-lyase (PAL) (Mizutani et al., 1995; Mizutani et al., 1997). In this study, I found that several P450 genes were also induced by light (Fig. 7B). It has been reported that CYP73A5, flavonoid 3'-hydroxylase (CYP73A5), was induced by light (Mizutani et al., 1995; Mizutani et al., 1997). Interestingly, CYP76C1, which was most strongly induced by light, was predominantly

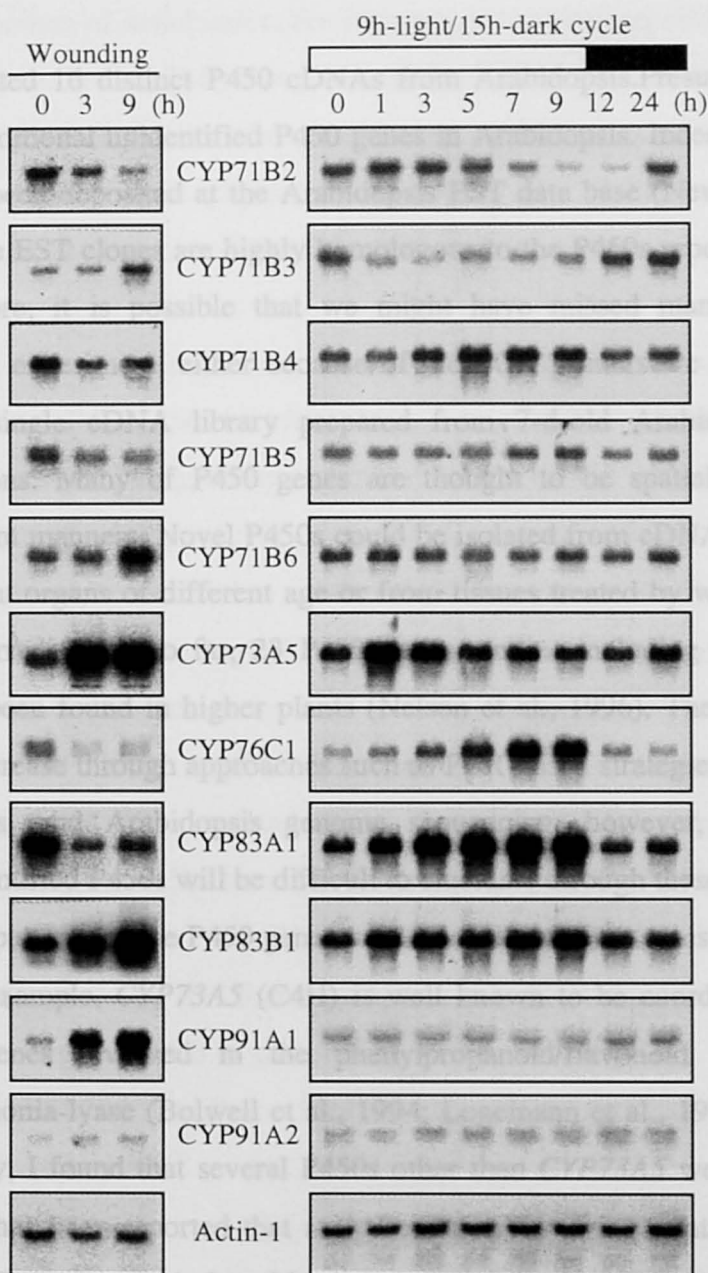


Figure 7. Effect of wounding and light/dark cycle on expression levels of the 11 P450 genes.

(A) Leaves were harvested from 3-week-old plants grown under continuous light. The harvested samples were cut into strips 2 mm in width, and incubated for 9h under continuous light in a petri dish containing GM medium. Total RNA was isolated at the times indicated after wounding and analysed by RNA gel blotting (5 µg per lane) using the probes indicated.

(B) Plants were grown under a 9h light/15h dark cycle: light period, 0-9h; dark period, 9-24h. Total RNA was isolated from the leaves of 3-week-old plants at times indicated after the onset of light period (0h), and analysed by RNA gel blotting (5 µg per lane) using the probes indicated.

Discussion

I have isolated 16 distinct P450 cDNAs from Arabidopsis. Presumably, there are a large number of additional unidentified P450 genes in Arabidopsis. Indeed, many unknown P450 clones have been deposited at the Arabidopsis EST data base (Newman et al., 1994), and several of these EST clones are highly homologous to the P450s reported here (data not shown). Furthermore, it is possible that we might have missed many P450 clones in Arabidopsis in this experiment, either because of the PCR primers we used, or because I screened only a single cDNA library prepared from 7-d-old Arabidopsis under high stringency conditions. Many of P450 genes are thought to be spatially and temporally regulated in different manners. Novel P450s could be isolated from cDNA libraries prepared from either different organs of different age or from tissues treated by wounding, pathogen infection, or UV irradiation. So far, 23 P450 gene families including those 10 found in Arabidopsis have been found in higher plants (Nelson et al., 1996). The number of cloned P450 genes will increase through approaches such as PCR-based strategies, Arabidopsis EST data base searches, and Arabidopsis genome sequencing; however, the physiological functions of the identified P450s will be difficult to elucidate through these efforts.

Expression patterns of the P450 genes *in planta* will in some cases have physiological implications. For example, *CYP73A5* (C4H) is well known to be coordinately induced by light with the genes involved in the phenylpropanoid/flavonoid pathway such as phenylalanine ammonia-lyase (Bolwell et al., 1994; Logemann et al., 1995; Mizutani et al., 1997). In this study, I found that several P450s other than *CYP73A5* were also induced by light (Fig. 7B). It has been reported that a pigmentation/flowering related P450, flavonoid 3'-hydroxylase (*CYP75*), was induced by light (Holton et al., 1993; Bolwell et al., 1994). Interestingly, *CYP76C1*, which was most strongly induced by light, was predominantly expressed in flowers (Fig. 7B).

Wounding treatment also resulted in strong induction of specific P450 species, while some of them were significantly downregulated (Fig. 7A). Namely, *CYP71B3*, *CYP71B6*, *CYP83B1*, *CYP91A1* and *CYP91A2* expression levels were higher in older leaves than in younger leaves, and these P450s were induced by wounding. It has been reported that *CYP71A1* was induced during increased ethylene biosynthesis triggered by wounding (Bolwell et al., 1994). P450s are also involved in defense mechanisms in plants, which include wound-healing, protection against UV irradiation, resistance against pathogen

infection and metabolism of xenobiotics. For instance, it has been reported that the levels of several P450s: C4H, isoflavone 2'-hydroxylase, and isoflavone synthase in alfalfa suspension cells and 6a-hydroxylase in soybean were enhanced by elicitor challenge, leading to the biosynthesis of the isoflavonoid phytoalexins (Bolwell et al., 1994). It remains to be investigated how these newly isolated P450s in *Arabidopsis* will respond under various stress conditions.

Novel P450 clones are interesting candidates for heterologous expression, and the expressed recombinant P450s can be characterized in terms of their catalytic activities. However, even if a reaction catalyzed by such a recombinant P450 can be identified, it is not necessarily a principal physiological role of the P450. In other words, it is possible that the identified reaction should be ascribed to another P450 *in vivo*, and the P450 of interest may have a completely different physiological function. An interesting approach is to use P450s of unknown functions to generate sense- and antisense-expressing transgenic plants. Also, genetic linkage may be used to explore functions of new P450s. Recently, several P450 clones have been isolated from mutated plants by either T-DNA tagging or transposon tagging, and functions of the P450s have been investigated (Winkler RG et al., 1995; Meyer K et al., 1996; Szekeres M et al, 1996; Bishop GJ et al, 1996). In this chapter, I determined the loci of four P450 genes by RFLP mapping, and it appears that one of them might be involved in a metabolic step related to gibberellins.

P450s of identified catalytic function represent promising candidates for manipulating specific plant traits.

Summary

I have isolated multiple cytochrome P450 (P450) cDNAs from *Arabidopsis thaliana* employing a PCR strategy. Degenerate oligonucleotide primers were designed from conserved amino acid sequences between two plant P450s, CYP71A1 and CYP73A2, including the heme-binding site and the proline rich motif found in the N-terminal region, and 11 putative P450 fragments were amplified from first strand cDNA from 7-d-old *Arabidopsis* as a template. With these PCR fragments as hybridization probes, 13 full-length and 3 partial cDNAs encoding different P450s have been isolated from an *Arabidopsis* cDNA library. These P450s have been assigned to either one of the established subfamilies:

CYP71B, CYP73A, and CYP83A; or novel subfamilies: CYP76C, CYP83B, and CYP91A. The primary protein structures predicted from the cDNA sequences revealed that the regions around both the heme-binding site and the proline-rich motif were highly conserved among all these P450s. The N-terminal structures of the predicted P450 proteins suggested that these *Arabidopsis* P450s were located at the endoplasmic reticulum membrane. The loci of four P450 genes were determined by RFLP mapping. One of the clones, *CYP71B2*, was located at a position very close to the *ga4* and *gai* mutations. RNA blot analysis showed expression patterns unique to each of the P450s in terms of tissue specificity and responsiveness to wounding and light/dark cycle, implicating involvement of these P450s in diverse metabolic processes.

Conclusion

Plant cytochrome P450 monooxygenase system is involved in a wide variety of metabolic processes of endogenous and exogenous lipophilic compounds. However, only a P450 cDNA with an unknown function had been isolated from avocado (Bozak et al., 1990) and only a few of plant P450s had been purified to homogeneity, when I started the investigation of plant cytochromes P450 in 1991.

In this thesis, I investigated stress responses of plant P450 monooxygenase system by focusing on a P450 possessing cinnamate 4-hydroxylase activity as well as NADPH-cytochrome P450 reductase. I also demonstrated diversity of P450 monooxygenase system in plants by isolation of multiple P450 cDNAs from a single plant species.

In Chapter I, I demonstrated the purification of the P450_{C4H} from mung bean seedlings and confirmed the C4H activity of the purified P450 in the reconstitution system. This provides the direct evidence that a physiological C4H activity is associated with a P450 protein. Then, I isolated the P450 cDNAs with the aid of partial amino acid sequences determined for the purified P450 (Chapter II). This is the first isolation of cDNA clones encoding a higher plant P450 possessing a clear physiological activity.

In Chapter III, I isolated a cDNA and a genomic clone encoding C4H from *Arabidopsis* in order to characterize a mechanism(s) by which transcription of the P450 is regulated in plants. I showed that the cDNA actually functions as a C4H protein by expressing a recombinant protein in insect cells and reconstituting its C4H activity with P450-reductase. The expression of the P450 gene was coordinated with other genes in phenylpropanoid pathway in response to light and wounding. I also analyzed the DNA sequence of the *C4H* promoter and found that the promoter contains all of the three *cis*-acting elements conserved among *PAL* and *4CL* genes so far reported. This is the first characterization of the P450 gene promoter in plants.

In Chapter IV, I isolated cDNA and genomic clones for two isoforms of P450-reductase from *Arabidopsis* in order to characterize the electron transfer system of plant P450 monooxygenases. The recombinant proteins for the two isoforms showed similar biochemical properties, including the reconstituted C4H activity. The two isoforms, on the other hand, showed the different expression patterns in tissue specificity and responsiveness to wounding

and light. The promoter of the inducible reductase contained the consensus sequence motifs involved in the coordinated regulation of phenylpropanoid pathway genes but the motifs were not found in the promoter of the constitutive reductase. These results suggest different mechanisms for their transcriptional activation and also their different physiological roles *in planta*.

In Chapter V, diversity of P450 genes in plants were explored by using molecular biological approaches. Sixteen different P450 cDNAs were isolated from Arabidopsis, and their deduced protein structures suggested the localization at ER membrane. As a first step towards the elucidation of their physiological functions, the loci of four P450 genes were determined by RFLP mapping. The isolated multiple P450 genes showed expression patterns unique to each of the P450s in terms of tissue specificity and stress responses such as wounding and light, implicating involvement of these P450s in diverse metabolic processes.

Recent efforts to explore novel P450s in plants revealed the great diversity of plant P450s. Since 1990, to date more than 100 P450s (!) from a variety of plant species have been sequenced and 30 P450 gene families exist in plants (Nelson et al, 1996; <http://drnelson.utmem.edu/homepage>). Many hydroxylation reactions in various biosynthetic pathway have been defined as P450-dependent reactions by biochemical approaches. However, there is still a great gap between biochemical and molecular biological studies of plant P450s, and physiological functions and substrate specificities of the isolated P450s must be investigated in the future. Heterologous expression of both P450s and P450-reductase and reconstitution of the P450 activities (as described in Chapter III, IV) will be one of the best approaches to overcome this problem. Molecular mechanisms by which P450 activities are regulated *in planta* also have to be studied. In this thesis, I investigated the transcriptional regulation of P450s and P450-reductases in response to stresses (Chapter III, IV, V). The results suggested that the P450 monooxygenase system does not have a regulation mechanism unique to the P450 system, but is coordinately regulated by a mechanism(s) similar to those conserved in the other enzymes which are physiologically linked to the P450 system.

References

- Appert C, Logemann E, Hahlbrock K, Schmid J, Amrhein N (1994) Structural and catalytic properties of the four phenylalanine ammonia-lyase isozymes from parsley (*Petroselinum crispum* Nym.). *Eur J Biochem* 225: 491-499.
- Ausubel FM, Brent R, Kingston RE, Moore DD, Siedman JG, Smith JA, Struhl K (1987) *Current Protocols in Molecular Biology*. John Wiley, New York
- Beltzer JP, Fiedler K, Fuhrer C, Geffen I, Handschin C, Wessels HP, Spiess M (1991) Charged residues are major determinants of the transmembrane orientation of a signal-anchor sequence. *J Biol Chem* 266: 973-978
- Benveniste I, Salaun JP, Durst F (1977) *Phytochemistry* 16: 69-73
- Benveniste I, Lesot A, Hasenfratz MP, Durst F (1989) Immunocharacterization of NADPH-cytochrome P450 reductase from Jerusalem artichoke and other higher plants. *Biochem J* 259: 847-853
- Benveniste I, Lesot A, Hasenfratz MP, Kochs G, Durst F (1991) Multiple forms of NADPH-cytochrome P450 reductase in higher plants. *Biochem Biophys Res Comm* 177(1): 105-112
- Bishop GJ, Harrison K, Jones JDG (1996) The tomato *Dwarf* gene isolated by heterologous transposon tagging encodes the first member of a new cytochrome P450 family. *Plant Cell* 8: 959-969
- Bolwell GP, Bozak K, Zimmerlin A (1994) Plant cytochrome P450. *Phytochemistry* 37(6): 1491-1506
- Bordier C (1981) Phase-separation of integral membrane proteins in Triton X-114 solution. *J Biol Chem* 256: 1604-1607
- Bozak KR, Yu H, Sirevag R, Chistoffersen RE (1990) Sequence analysis of ripe-related cytochrome P450 cDNAs from avocado fruits. *Proc Natl Acad Sci USA* 87: 3904-3908
- Cabib E, Polacheck I (1984) Protein assay for dilute solutions. *Methods Enzymol* 104:415-416
- Chang C, Bowman JL, DeJohn AW, Lander ES, Meyerowitz EM (1988) Restriction fragment length polymorphism linkage map for *Arabidopsis thaliana*. *Proc Natl Acad Sci USA* 85: 6856-6860
- Chiang H-H, Hwang I, Goodman HM (1995) Isolation of the Arabidopsis GA4 locus. *Plant Cell* 7: 195-201
- Chapple CCS (1995): A cDNA encoding a novel cytochrome P450-dependent monooxygenase from Arabidopsis thaliana. *Plant Physiol* 108: 875-876
- Church GM, Gilbert W (1984) Gemonic sequencing. *Proc Natl Acad Sci USA* 81: 1991-1995
- Chappell J, Hahlbrock K (1984) Transcriptional activation of plant defense genes in response to UV-light or fungal elicitor. *Nature* 311: 76-78
- Cuatrecasas O (1970) Protein purification by affinity chromatography. Derivatization of agarose and polyacrylamide beads. *J Biol Chem* 245: 3059-3065
- Dangl JL (1992) Regulatory elements controlling developmental and stress-induced expression of phenylpropanoid genes. In "Plant Gene Research", Boller T and Meins F, eds. Vol. 8, Springer-Verlag/Wien, New York, pp 303-326
- Devereux J, Haeberli P, Smithies O (1984) A comprehensive set of sequence analysis programs for the VAX. *Nucl Acids Res* 12: 387-395
- Dixon RA, Harrison MJ (1990) Activation, structure, and organization of genes involved in microbial defense in plants. *Adv Genet* 28: 165-234

- Dixon RA, Paiva NL (1995) Stress-induced phenylpropanoid metabolism. *Plant Cell* 7: 1085-1097
- Donaldson RP, Luster DG (1991) Multiple forms of plant cytochromes P450. *Plant Physiol* 96: 669-674
- Douglas CJ (1996) Phenylpropanoid metabolism and lignin biosynthesis: from weeds to trees. *Trends in Plant Science* 6: 171-178
- Douglas CJ, Hoffmann H, Schulz W, Hahlbrock K (1987) Structure and elicitor or UV-light stimulated expression of two 4-coumarate: CoA ligase genes in parsley. *EMBO J* 6: 1189-1195
- Fahrendorf T, Dixon RA (1993) Stress responses in alfalfa (*Medicago sativa* L.) XVIII: Molecular cloning and expression of the elicitor-inducible cinnamic acid 4-hydroxylase cytochrome P450. *Archives Biochem Biophys* 305: 509-515
- Feinbaum RL, Ausubel FM (1988) Transcriptional regulation of the *Arabidopsis thaliana* chalcone synthase gene. *Molecular and Cellular Biology* 8: 1985-1992
- Feinberg AP, Vogelstein B (1983) A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal Biochem* 152: 232-238
- Fonne-Pfister R, Kreuz K (1990) Ring-methyl hydroxylation of chlorotoluron by an inducible cytochrome P450-dependent enzyme from maize. *Phytochemistry* 29: 2793-2796
- Frank MR, Deyneka JM, Schuler MA (1996) Cloning of wound-inducible cytochrome P450 monooxygenases expressed in pea. *Plant Physiol* 110: 1035-1046
- Frey M, Kliem R, Saedler H, Gierl A (1995) Expression of a cytochrome P450 gene family in maize. *Mol Gen Genet* 246:100-109
- Gabriac B, Werck-Reichhart D, Teusch H, Durst F (1991) Purification and immunocharacterization of a plant cytochrome P450: The cinnamate 4-hydroxylase. *Arch Biochem Biophys* 288: 302-309
- Gotoh O, Fujii-Kuriyama Y (1989) Evolution, structure, and gene regulation of cytochrome P450. In "Basis and Mechanisms of Regulation of Cytochrome P450". Ruckpaul K and Rein H, eds. Vol.1, Taylor and Francis, London, New York and Philadelphia, pp195-243
- Graebe JE (1987) Gibberellin biosynthesis and control. *Ann Rev Plant Physiol* 38: 419-65
- Grossberger D (1987) Minipreps of DNA from bacteriophage lambda. *Nucleic Acids Res* 15: 6737
- Hahlbrock K, Scheel D (1989) Physiology and molecular biology of phenylpropanoid metabolism. *Ann Rev Plant Physiol Plant Mol Biol* 40: 347-369
- Hallahan DL, Nugent, JHA, Hallahan BJ, Dawson GW, Smiley DW, West JM, Wallsgrove RM (1992) Interaction of avocado (*Persea americana*) cytochrome P450 with monoterpenoids. *Plant Physiol* 98: 1290-1297
- Hamerski D, Schmitt D, Matern U (1990) Induction of two prenyltransferases for the accumulation of coumarin phytoalexins in elicitor-treated *Ammi majus* cell suspension cultures. *Phytochemistry* 29: 1131-1135
- Hanley BA, Schuler MA (1988) Plant intron sequences: evidence for distinct groups of introns. *Nucl Acid Res* 16: 7159-7174
- Hattori M, Sakaki Y (1986) Dideoxy sequencing method using denatured plasmid templates. *Anal Biochem* 162: 232-238
- Hatzios KK (1991) Biotransformations of herbicides in higher plants. In *Environmental Chemistry of Herbicides*, Vol. 2. Edited by Grover R and Cessna AJ: 141-185 CRC Press, Boston.

- Hauffe KD, Lee SP, Subramaniam R, Douglas CJ (1993) Combinatorial interactions between positive and negative cis-acting elements control spatial patterns of *4CL-1* expression in transgenic tobacco. *Plant J* 4: 235-253
- Hauge BM, Hanley SM, Cartinhour S, Cherrt JM, Goodman HM, Koornneef M, Stam P, Chang C, Kempin S, Medrano L, Meyerowitz EM (1993) An integrated genetic/RFLP map of the *Arabidopsis thaliana* genome. *Plant J* 3: 745-754
- Higashi K, Ikeuchi K, Obara M, Karasaki Y, Hirano H, Gotoh S, Koga Y (1985) Purification of a single major form of microsomal cytochrome P450 from tulip bulbs (*Tulipa gesneriana* L.). *Agric Biol Chem* 49: 2399-2405
- Holton TA, Brugliera F, Lester DR, Tanaka Y, Hyland CD, Menting JGT, Lu CY, Farcy E, Stevenson TW and Cornish EC (1993) Cloning and expression of cytochrome P450 genes controlling flower color. *Nature* 366: 276-279
- Hotze M, Schroder G, Schroder J (1995) Cinnamate 4-hydroxylase from *Catharanthus roseus*, and a strategy for the functional expression of plant cytochrome P450 proteins as translational fusions with P450 reductase in *Escherichia coli*. *FEBS Lett* 374: 345-350
- Imai Y (1976) The use of 8-aminooctyl Sepharose for separation of some components of the hepatic microsomal electron transfer system. *J Biochem* 80: 267-276
- Iyanagi T, Mason HS (1973) Some properties of hepatic reduced nicotinate adenine dinucleotide phosphate-cytochrome c reductase. *Biochemistry* 12: 2297-2308
- Johnson EF (1992) Mapping determinants of the substrate selectivities of P450 enzymes by site-directed mutagenesis. *Trends Pharmacol Sci* 13: 122-126
- Kozak M (1987) An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. *Nucl Acid Res* 15: 8125-8148
- Kochs G, Griesbach H (1989) Phytoalexin synthesis in soybean: purification and reconstitution of cytochrome P450 3,9-dihydroxypterocarpan 6 α -hydroxylase and separation from cytochrome P450 cinnamate 4-hydroxylase. *Arch Biochem Biophys* 273: 543-553
- Koch G, Werck-Reichhart D, Griesbach H (1992) Further characterization of cytochrome P450 involved in phytoalexin synthesis in soybean: cytochrome P450 cinnamate 4-hydroxylase and 3,9-dihydroxypterocarpan 6 α -hydroxylase. *Arch Biochem Biophys* 293: 187-194
- Koch BM, Sibbesen O, Halkier BA, Svendsen I, Moller BL (1995) The primary sequence of cytochrome P450_{tyr}, the multifunctional N-hydroxylase catalyzing the conversion of L-tyrosine to p-hydroxyphenylacetaldehyde oxime in the biosynthesis of cyanogenic glucoside dhurrin in *Sorghum bicolor* (L.) Moench. *Arch Biochem Biophys* 323: 177-186
- Kraus PFX and Kutchan TM (1995) Molecular cloning and heterologous expression of a cDNA encoding berbamine synthase, a C-O phenol-coupling cytochrome P450 from the higher plant *Berberis stolonifera*. *Proc Natl Acad Sci USA* 92: 2071-2075
- Laemmli UK (1970) Cleavage of structural proteins during assembly of the head of bacteriophage T4. *Nature* 227: 680-685
- Lagrimini LM, Burkhardt W, Moyer M, Rothstein S (1987) Molecular cloning of complementary DNA encoding the lignin-forming peroxidase from tobacco: Molecular analysis and tissue-specific expression. *Proc Natl Acad Sci USA* 84: 7542-7546
- Lamb CJ (1977) *trans*-Cinnamic acid as a mediator of the light stimulated increase in hydroxycinnamoyl-CoA:quinate hydroxycinnamoyl transferase. *FEBS Lett* 75: 37-40

- Lander LRR, Green MP, Abrahamson J, Barlow A, Dale M, Lincoln SE, Newsburg L (1987) MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. *Genomics* 1: 174-181
- Lawton MA, Lamb CJ (1987) Transcriptional activation of plant defense genes by elicitor, wounding, and infection. *Mol Cell Biol* 7: 335-341
- Lee D, Ellard M, Wanner LA, Davis KR, Douglas CJ (1995) The *Arabidopsis thaliana* 4-coumarate:CoA ligase(4CL) gene: stress and developmentally regulated expression and nucleotide sequence of its cDNA. *Plant Mol Biol* 28: 871-884
- Leon J, Shulaev V, Yalpani N, Lawton MA, Raskin I (1995) Benzoic acid 2-hydroxylase, a soluble oxygenase from tobacco, catalyzes salicylic acid biosynthesis. *Proc Natl Acad Sci USA* 92: 10413-10417
- Lindberg RL, Negishi M (1989) Alteration of mouse cytochrome P450c₁₇ substrate specificity by mutation of a single amino-acid residue. *Nature* 339: 632-634
- Logemann E, Parniske M, Hahlbrock K (1995) Modes of expression and common structural features of the complete phenylalanine ammonia-lyase gene family in parsley. *Proc Natl Acad Sci USA* 92: 5905-5909
- Lois R, Dietrich A, Hahlbrock K, Schulz W (1989) A phenylalanine ammonia-lyase gene from parsley: structure, regulation and identification of elicitor and light responsive cis-acting elements. *EMBO J* 8: 1641-1648
- Lois R, Hahlbrock K (1992) Differential wounding activation of members of the phenylalanine ammonia-lyase and 4-coumarate: CoA ligase gene families in various organs of parsley plants. *Z Naturforsch* 47c: 90-94
- Lowry OH, Rosebrough NJ, Farr AL, Randall RJ (1951) Protein measurement with the Folin reagent. *J Biol Chem* 193: 265-275
- Madyastha KM, Ridgway JE, Dwyer JG, Coscia CJ (1977) Subcellular localization of a cytochrome P450 dependent monooxygenase in vesicles of the higher plant *Catharanthus roseus*. *J Cell Biol* 72: 302-313
- Markham A, Hartman GC, Parke DV (1972) *Biochem J* 130: 90P
- Masters, BSS, Okita RT (1982) The history, properties and function of NADPH-cytochrome P450 reductase. In *Hepatic Cytochrome P450 Monooxygenase system*, edited by Schenkman JB, Kupfer D (Oxford: Pergamon Press), pp. 342-360
- Maüch-Mani B, Slusarenko AJ (1996) Production of Salicylic acid precursors in a major function of phenylalanine ammonia-lyase in the resistance of *Arabidopsis* to *Peronospora parasitica*. *Plant Cell* 8: 203-212
- Meijer AH, Cardoso MIL, Voskuilen JT, de Waal A, Verpoorte R, Hoge JHC (1993) Isolation and characterization of a cDNA clone from *Catharanthus roseus* encoding NADPH-cytochrome P450 reductase, an enzyme essential for reactions catalyzed by cytochrome P450 monooxygenases in plants. *Plant J* 4(1): 47-60
- Meijer AH, Souer E, Verpoorte R, Hoge JHC (1993) Isolation of cytochrome P450 cDNA clones from the higher plant *Catharanthus roseus* by a PCR strategy. *Plant Mol Biol* 22: 379-383
- Meyer K, Cusumano JC, Somerville C, Chapple CCS (1996) Ferulate 5-hydroxylase from *Arabidopsis thaliana* defines a new family of cytochrome P450-dependent monooxygenases. *Proc Natl Acad Sci USA* 93: 6869-6874
- Mizutani M, Ohta D, Sato R (1993a) Purification and characterization of a cytochrome P450 (*trans*-cinnamic acid 4-hydroxylase) from etiolated mung bean seedlings. *Plant Cell Physiol* 34: 481-488
- Mizutani M, Ward E, DiMaio J, Ohta D, Ryals J, Sato R (1993b) Molecular cloning and sequencing of a cDNA encoding mung bean cytochrome P450(P450C4H) possessing cinnamate 4-hydroxylase activity. *Biochem Biophys Res Comm* 190(3): 875-880

Mizutani M, Ohta D, Sato R (1997) Isolation of a cDNA and a genomic clone encoding cinnamate 4-hydroxylase from *Arabidopsis thaliana* and its expression manner in *planta*. *Plant Physiol* 113: 755-763

Mol JMN, Stuitje AR, Gerats AGM, Koes RE (1988) Cloned genes of phenylpropanoid metabolism in plants. *Plant Mol Biol Rep* 6: 274-278

Nairn CJ, Winsett L, Ferl RJ (1988) Nucleotide sequence of an actin gene from *Arabidopsis thaliana*. *Gene* 65: 247-257

Negrak V, Yang P, Subramanian M, McNevin JP, Lemieux B (1996) Molecular cloning and characterization of the *CER2* gene of *Arabidopsis thaliana*. *Plant J* 9: 137-145

Nelson DR, Kamataki T, Waxman DJ, Guengerich FP, Estabrook RW, Feyereisen R, Gonzalez FJ, Coon MJ, Gunsalus IC, Gotoh O, Okuda K, and Nebert DW (1993) The P450 superfamily: update on new sequences, gene mapping, accession numbers, early trivial names of enzymes, and nomenclature. *DNA Cell Biol* 12: 1-51

Nelson DR, Koymans L, Kamataki T, Stegeman JJ, Feyereisen R, Waxman DJ, Waterman MR, Gotoh O, Coon MJ, Estabrook RW, Gunsalus IC and Nebert DW (1996) P450 superfamily: update on new sequences, gene mapping, accession numbers and nomenclature. *Pharmacogenetics* 6: 1-42

Nelson DR, Strobel HW (1988) On the membrane topology of vertebrate cytochrome P450 proteins. *J Biol Chem* 263: 6038-6050

Newman T, de Bruijn FJ, Green P, Keegstra K, Kende H, McIntosh L, Ohlrogge J, Raikhel N, Somerville S, Thomashow M, Retzel E, Somerville C (1994) Gene Galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous *Arabidopsis* cDNA clones. *Plant Physiol* 106: 1241-1255

Nishimoto Y (1986) Localization of cytochrome c -binding domain on NADPH-cytochromeP450 reductase. *J Biol Chem* 261: 14232-14239

Oba K, Conn EE (1988) Induction of cinnamate 4-hydroxylase in developing maize seedlings. *Phytochemistry* 27: 2447-2450

O'Keefe DP, Leto KJ (1989) Cytochrome P450 from the mesocarp of avocado (*Persea americana*). *Plant Physiol* 89: 1141-1149

O'Leary KA, Beck TW, Kasper CB (1994) NADPH-cytochrome P450 oxidoreductase gene: identification and characterization of the promoter region. *Arch Biochem Biophys* 310(2): 452-459

Ohl S, Hedrich S, Chory J, Lamb CJ (1990) Functional properties of phenylalanine ammonia-lyase promoter from *Arabidopsis*. *Plant Cell* 2: 837-848

Omura T, Sato R (1964) The carbon monoxide binding pigment of liver microsomes. I. Evidence for its hemoprotein nature. *J Biol Chem* 239: 2370-2378

Pan Z, Durst F, Werck-Reichhart DW, Gardner HW, Camara B, Cournish K, Backhaus RA (1995) The major protein of guayule rubber particles is a cytochrome P450. *J Biol Chem* 270(15): 8487-8494

Payne G, Ward E, Gaffney T, Ahl-Goy P, Moyer M, Harper A, Meins F, Ryals J (1990) Evidence for a third structural class of β -1,3-glucanase in tobacco. *Plant Mol Biol* 15: 797-808

Persans MW, Schuler MA (1995) Differential induction of cytochrome P450-mediated triasufuron metabolism by naphthalic anhydride and triasulfuron. *Plant Physiol* 109: 1483-1490

Porter TD, kasper CB (1986) NADPH-cytochrome P450 oxidoreductase: flavin mononucleotide and flavin adenine dinucleotide domains evolved from different flavoproteins. *Biochemistry* 25: 1682-1687

Porter TD, Thomas WB, Kasper CB (1990) NADPH-cytochrome P450 oxidoreductase gene organization correlates with structural domains of the protein. *Biochemistry* 29: 9814-9818

Porter TD (1991) An unusual yet strongly conserved flavoprotein reductase in bacteria and mammals. *Trends Biochem Sci* 16: 154-158

Potter S, Moreland DE, Kreuz K, Ward E (1995) Induction of cytochrome P450 genes by ethanol in maize. *Drug Metab Drug Interact* 12: 317-327

Potts JRM, Weklych R, Conn EE (1974) The 4-hydroxylation of cinnamic acid by Sorghum microsomes and the requirement for cytochrome P450. *J Biol Chem* 249: 5019-5026

Reichhart D, Salaun JP, Benveniste I, Durst F (1980) Time course of induction of cytochrome P450, NADPH-cytochrome c reductase, and cinnamic acid 4-hydroxylase by phenobarbital, ethanol, herbicides, and manganese in higher plant microsomes. *Plant Physiol* 66: 600-604

Rich PR, Bendall DS (1975) Cytochrome components of plant microsomes. *Eur J Biochem* 55: 333-341

Riviere JL and Cabbane F (1987) Animal and plant cytochrome P450 systems. *Biochimie* 69: 743-752

Ruckpaul K, Rein H, Blanck J (1989) Regulation mechanisms of the activity of the hepatic endoplasmic cytochrome P450. In "Basis and Mechanisms of Regulation of Cytochrome P450". Ruckpaul K and Rein H, eds. Vol.1, Taylor and Francis, London, New York and Philadelphia, pp 1-65

Russel DW (1971) The metabolism of aromatic compounds in higher plants. X. Properties of the cinnamic acid 4-hydroxylase of pea seedlings and some aspects of its metabolic and developmental control. *J Biol Chem* 246: 3870-3878

Sandermann H (1992) Xenobiotics. *Trends Biochem Sci* 17: 82-84

Shet MS, Sathasivan K, Arlotto MA, Mehdy MC, Estabrook RW (1993) Purification, characterization, and cDNA cloning of an NADPH-cytochrome P450 reductase from mung bean. *Proc Natl Acad Sci USA* 90: 2890-2894

Song WC, Brash AR (1991) Purification of an allene oxide synthase and identification of the enzyme as a cytochrome P450. *Science* 253: 781-784

Song WC, Funk CD, Brash AR (1993) Molecular cloning of an allene oxide synthase: A cytochrome P450 specialized for the metabolism of fatty acid hydroperoxides. *Proc Natl Acad Sci USA* 90: 8519-8523

Spatz L, Strittmatter P (1971) A form of cytochrome b5 that contains an additional hydrophobic sequence of 40 amino acid residues. *Proc Natl Acad Sci USA* 68: 1042-1046

Steffens M, Ettl F, Kranz D, Kindl H (1989) Vanadate mimics effects of fungal cell wall in eliciting gene activation in plant cell cultures. *Planta* 177: 160-168

Summers MD, Smith GE (1987) In "A Manual of Methods for Baculovirus Vectors and Insect Cell Culture Procedures", bulletin No.1555. Texas Agricultural Experiment Station and Texas A and M University, College Station, TX

Szczena-Skorupa E, Browne N, Mead D, Kemper B (1988) Positive charges at the NH2 terminus convert the membrane-anchor signal peptide of cytochrome P450 to a secretory signal peptide. *Proc Natl Acad Sci USA* 85: 738-742

Szczesna-Skorupa E, Straub P, Kemper B (1993) Deletion of a conserved tetrapeptide, PPGP, in P4502C2 results in loss of enzymatic activity without a change in its cellular location. *Arch Biochem Biophys* 304: 170-5

Szekeres M, Nemeth K, Koncz-Kalman Z, Mathur J, Kauschmann A, Altmann T, Redei GP, Nagy F, Schell J, Koncz C (1996) Brassinosteroids rescue the deficiency of CYP90, a cytochrome P450, controlling cell elongation and de-etiolation in *Arabidopsis*. *Cell* 85: 171-182

Teutsch HG, Hasenfratz MP, Lesot A, Stoltz C, Garnier JM, Jeltsch JM, Durst F, Werck-Reichhart D (1993) Isolation and sequence of a cDNA encoding the Jerusalem artichoke cinnamate 4-hydroxylase, a major plant cytochrome P450 involved in the general phenylpropanoid pathway. *Proc Natl Acad Sci USA* 90: 4102-4106

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22: 4673-4680

Toguri T, Kobayashi O, Umemoto N (1993) The cloning of eggplant seedling cDNAs encoding proteins from a novel cytochrome P450 family (CYP76). *Biochem Biophys Acta* 1216: 165-169

Umemoto N, Kobayashi O, Ishizaki-Nishizawa O, Toguri T (1993) cDNAs sequences encoding cytochrome P450 (CYP71 family) from eggplant seedlings. *FEBS* 330(2): 169-173

Valvekens D, Montagu MV, Lusebettens MV (1988) *Agrobacterium tumefaciens*-mediated transformation of *Arabidopsis thaliana* root explants by using kanamycin selection. *Proc Natl Acad Sci USA* 85: 5536-5540

Vetter HP, Mangold U, Schroder G, Marner FJ, Werck-Reihhart D, Schroder J (1992) Molecular analysis and heterologous expression of an inducible cytochrome P450 protein from periwinkle (*Catharanthus roseus* L.). *Plant Physiol* 100: 998-1007

Wanner LA, Li G, Ware D, Somssich IE, Davis KR (1995) The phenylalanine ammonia-lyase gene family in *Arabidopsis thaliana*. *Plant Mol Biol* 27: 327-338

Werck-Reichhart D, Benveniste I, Teusch H, Durst F, Gabriac B (1991) Glycerol allows low temperature phase-separation of membrane proteins solubilized in Triton X-114: application to the purification of plant cytochromes P450 and b5. *Anal Biochem* 197: 125-131

West CA (1980) Hydroxylase, monooxygenase, and cytochrome P450. In *Biochemistry in plants*, Vol. 2. Edited by Davis DD pp. 317-364. Academic Press, New York

Winkler RG, Helentjaris T (1995) The maize Dwarf3 gene encodes a cytochrome P450-mediated early step in gibberellin biosynthesis. *Plant Cell* 7: 1307-1317

Yabusaki Y, Murakami H, Ohkawa H (1988) Primary structure of *Saccharomyces cerevisiae* NADPH-cytochrome P450 reductase deduced from nucleotide sequence of its cloned gene. *J Biochem* 103: 1004-1010

Yamazaki S, Sato K, Suhara K, Sakaguchi M, Mihara K, Omura T (1993) Importance of the proline-rich region following signal-anchor sequence in the formation of correct conformation of microsomal cytochrome P450s. *J biochem* 114: 652-7

Yalpani N, Leon J, Lawton MA, Raskin I (1993) Pathway of salicylic acid biosynthesis in healthy and virus-inoculated tobacco. *Plant Physiol* 103: 315-321

Zeevaart JAD, Talon M (1991) Gibberellin mutants in *Arabidopsis thaliana*. In CM Karssen et al., eds, *Plant Growth Regulation*. Kluwer, Amsterdam

Zimmerlin A, Durst F (1990) Xenobiotic metabolism in plants: aryl hydroxylation of diclofop by a cytochrome P450 enzyme from wheat. *Phytochemistry* 29: 1729-1732

Zimmerlin A, Durst F (1992) Aryl hydroxylation of the herbicide diclofop by a wheat cytochrome P450 monooxygenase. *Plant Physiol* 100: 874-881

List of Publications

- Matoh T, Ishigaki K, **Mizutani M**, Matsunaga W, Takabe K (1992) Boron nutrition of cultured tobacco BY-2 cells. I. Requirement for and intracellular localization of boron and selection of cells that tolerate low levels of boron. *Plant Cell Physiol* 33: 1135-1141
- Matoh T, Matsunaga W, **Mizutani M**, Ohno K (1993) A sensitive and microscale method for boron determination of plant materials. *Jpn J Soil Sci Plant Nutr* 64: 71-73
- * **Mizutani M**, Ohta D, Sato R (1993) Purification and characterization of a cytochrome P450 (*trans*-cinnamic acid 4-hydroxylase) from etiolated mung bean seedlings. *Plant Cell Physiol* 34: 481-488
- * **Mizutani M**, Ward E, DiMaio J, Ohta D, Ryals J, Sato R (1993) Molecular cloning and sequencing of a cDNA encoding mung bean cytochrome P450 (P450C4H) possessing cinnamate 4-hydroxylase activity. *Biochem Biophys Res Comm* 190: 875-880
- * **Mizutani M**, Ohta D, Sato R (1997) Isolation of a cDNA and a genomic clone encoding cinnamate 4-hydroxylase from *Arabidopsis thaliana* and its expression manner *in planta*. *Plant Physiol* 113: 755-763
- * **Mizutani M**, Ohta D (1997) Two isoforms of NADPH-cytochrome P450 reductase of *Arabidopsis thaliana*: gene structure, heterologous expression in insect cells, and differential expression *in planta*. submitted
- * **Mizutani M**, Ward E, Ohta D (1997) Cytochrome P450 superfamily in *Arabidopsis thaliana*: isolation of cDNAs, differential expression, and RFLP mapping of multiple cytochromes P450. submitted
- * indicates the manuscripts which are presented in this thesis.

Acknowledgments

I wish to express my sincere gratitude to Professor Jiro Sekiya, Kyoto University, for his invaluable advice and continuous encouragement.

The present work has been accomplished at International Research Laboratories, Ciba-Geigy Japan under the direction of Emeritus Professor Ryo Sato, and Dr. Daisaku Ohta.

I wish to express my sincere gratitude to the late Emeritus Professor Ryo Sato, Osaka University, for his invaluable advice, discussion, and continuous encouragement.

I am deeply grateful to Dr. Daisaku Ohta for his guidance, valuable discussion, intimate encouragement, and critical review throughout this study.

I also wish to express my thanks to Dr. Eric Ward, and the other members of Agricultural Biotechnology Research Unit, Ciba-Geigy Corp. USA, for their valuable guidance and kind support during my stay at the laboratory.

Special thanks must be given to Ms. Nobuko Uodome, Ms. Keiko Nagai, and Ms. Emi Ikui for their excellent technical assistance.

I also thank to the members of International Research Laboratories, Ciba-Geigy Japan for their technical assistance and continuous encouragement.

Finally, I am much indebted to my parents and my brother for their heartfelt encouragement. Furthermore, I am eager to give my best appreciation to my wife Masako and my son Azuki for her all and his charm throughout my research.

Mizutani, Masaharu